AD 762 138

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

Army Research Office
Durham, North Carolina

September 1972

AD 762138

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

PART 1

D D C

JUN 15 1973

13

Approved for public release; distribution unlimited.
The findings in this report are not to be construed
as an official Department of the Army position, un-
less so designated by other authorized documents.

Sponsored by
The Army Mathematics Steering Committee
on Behalf of

THE OFFICE OF THE CHIEF OF RESEARCH AND DEVFLOPMENT

530

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.

PROCEEDINGS OF THE SEVENTEENTH CONFERENCE

ON THE DESIGN OF EXPERIMENTS

PART I

Sponsored by the Army Mathematics Steering Committee

HOST

Walter Reed Army Institute of Research

Walter Reed Army Medical Center

27-29 October 1971

FOREWARD

The Walter Reed Army Institute of Research (WRAIR), the host for the Seventeenth Conference on the Design of Experiments in Army Research, Development and Testing, has as its basic mission to provide the medical research and professional graduate training required by the Army to fulfill its role in National Defense. Since statistics and other scientific disciplines play an ever increasing role in medical fields, it is not surprising that Colonel Edward L. Buescher, Director and Commandant of WRAIR, was pleased to have his installation serve as host to this design of experiments conference. Colonel Hinton J. Baker was asked to act as Chairman on Local Arrangements. He was assisted in this capacity by Douglas Tang. Those in attendance are in debt to these two gentlemen for so abling handling the many housing and transportation problems that arose during the course of the meeting. Major General Colin F. Vorder Bruegge, Commanding General of the Walter Reed Army Medical Center, in his welcoming remarks, as well as in his comments at the banquet, made the audience feel that their scientific accomplishments were helping medical research to develop in many areas.

This is the second time that WRAIR has served as host to one of these conferences. The eighth conference in this series, sponsored by the Army Mathematics Steering Committee on behalf of the Office of the Chief of Research and Development, Department of the Army, was held 24-26 October 1962 at the Walter Reed Army Medical Center. Of the 180 attendees at the eighth conference, 21 attended the seventeenth conference. Dr. Stefano Vivona, now with the American Cancer Society, was the local chairman for the earlier conference. He was one of the repeat attendees. Dr. Herbert C. Batson was one of the invited speakers at the 1962 conference. At the banquet of this meeting Dr. Baker presented him a citation for his outstanding scientific contribution to the field of medicine. Drs. Marvin Zelen and George Lavin presented papers at both conferences. At the eighth Conference Zelen contributed a paper which was jointly authored by Dr. Badrig Kurkjian and himself. At the 1971 conference he was an invited speaker. Another invited speaker who attended both meetings was Professor Bernard Greenberg. It is interesting to note that, among the others who attended both meetings and participated in both programs as chairman, panelist, or by contributing papers were O. P. Bruno, A. C. Cohen, Francis Drassel, Henry Ellner, Walter D. Foster, Frank E. Grubbs, Boyd Harshbarger, Badrig Kurkjian, Clifford J. Maloney and Beatrice S. Orleans.

We note a few more statistics about the two meetings. The earlier one had forty-four more registered persons. One of the features of these meetings, the clinical session, has gained in popularity. In the 1962 conference there were two of these sessions with only three papers, while in 1971 there were five sessions in which eleven clinical--type papers were presented. The percentage increase in the papers for the technical sessions were not as great, still there was some increase, namely twenty-six contributed papers in 1962 as compared with thirty-two technical papers presented this year.

One of the requirements the chairman makes of the members of his Program Committee is that they name at least one speaker whose topic will be of special interest to the members of the host installation. This year that request was taken seriously. Most of the invited speakers touched on topics that were of interest to those in the medical field. This can be seen from the following list of invited speakers and the titles of their addresses:

"The Role of Mathematical Sciences in Biomedical Research" by Professor Marvin Zelen

iii

Preceding page blank

"Randomized Response: A New Survey Tool to Collect Data of a
Personal Nature" by Professor Bernard G. Greenberg

"Classification and Clustering Techniques in Data Analysis"
by Dr. Geoffrey H. Ball

"Hotelling's Weighing Desings"
by Professor K. S. Banerjee

"The Comparison of Proportions: A review of Significance Tests,
Confidence Intervals and Adjustments for Stratification"
by Dr. John J. Gart

Let us take this opportunity to thank members of the program committee
(David Alling, Hinton J. Baker, Francis Dressel, Walter Foster, Fred Frishman,
Bernard Greenberg, Bernard Harris, Boyd Harshbarger, Allyn Kimball, Clifford J.
Maloney, Herbert Solomon and Douglas Tang) for their recommendations for invited
speakers, as well as their active part in the conduction of this meeting. We
would be remiss in our duties if we did not give due praise to all those indi-
viduals who gave contributed papers. For truly, without their help, this meeting
could not have been a successful scientific conference.

Members of the Army Mathematics Steering Committee have asked that the
proceedings of this conference be made available to those in attendance for
further study of the contents of the presented papers, and to those interested
in the topics covered at the conference but who were unable to attend.

Francis G. Dressel   Frank E. Grubbs
Secretary   Conference Chairman

iv

TABLE OF CONTENTS

PART 1 and 2

vii

SEVENTEENTH CONFERENCE ON THE DESIGN OF EXPERIMENTS

IN ARMY RESEARCH, DEVELOPMENT AND TESTING

27-29 October 1971

Wednesday, 27 October

0830-0930   REGISTRATION - Lobby of Sternberg Auditorium (WRAIR)

0930-1220   GENERAL SESSION 1 - Sternberg Auditorium

Chairman: Colonel Lothrop Mittenthal, Office of the
Chief of Research and Development, Washington, DC

THE ROLE OF MATHEMATICAL SCIENCES IN BIOMEDICAL RESEARCH

Professor Marvin Zelen, State University of New York
at Buffalo, Amherst, New York

RANDOMIZED RESPONSE: A NEW SURVEY TOOL TO COLLECT DATA
OF A PERSONAL NATURE

Bernard G. Greenberg, The University of North Carolina,
Chapel Hill, North Carolina

1220-1320   LUNCH - Ballroom, Officers' Open Mess, WRAMC

1320-1510   CLINICAL SESSION A - Sternberg Auditorium

Chairman: Douglas Tang, Division of Biometrics
and Medical Information, Processing, Walter
Reed Army Institute of Research, Walter Reed
Army Medical Center, Washington, D. C.

METHODS TO EXTEND THE UTILITY OF LINEAR DISCRIMINANT
ANALYSIS

Captain L. E. Larsen, Division of Neuropsychiatry,
Walter Reed Army Institute of Research, Walter Reed
Army Medical Center, Washington, D. C.

PROBLEMS IN CHARACTERIZATION AND ANALYSIS OF PSYCHO-
PHYSIOLOGICAL RESPONSES

Captain John R. Jennings, Department of Experimental
Psychophysiology, Walter Reed Army Institute of Re-
search, Walter Reed Army Medical Center, Washington,
D. C.

1510-1540   Break - Lobby of Sternberg Auditorium

1540-1720   CLINICAL SESSION B - Sternberg Auditorium

Chairman: Badrig M. Kurkjian, US Army Materiel
   Command, Washington, D. C.

UTILIZATION OF GAS FLOW DYNAMICS IN A GRENADE LAUNCHER
SYSTEM

Herman E. Tarnow, US Army Weapons Command, Small
Arms Weapons System Directorate, Rock Island, Illinois

SIMULATING SUBSURFACE NUCLEAR EXPLOSIONS WITH CHEMICAL
EXPLOSIVES

Edward J. Leahy, Explosive Excavation Research Office,
Lawrence Livermore Laboratory, Livermore, California

1540-1720   TECHNICAL SESSION 3 - Room 358

Chairman: Beatrice S. Orleans, Naval Ships Systems
   Command, Washington, D.C.

MODIFIED FACTORIAL EXPERIMENTS FOR POISSON DATA

Lyman Ott and William Mendenhall, Department of
Statistics, University of Florida, Gainesville, Florida

INFERENCES OF FUNCTIONS OF THE PARAMETERS OF THE WEIBULL
DISTRIBUTION

Ronald L. Racicot, Applied Mathematics and Mechanics
Brance, Benet R&E Labs, Watervliet Arsenal, Watervliet,
New York

ESTIMATION IN THE EXPONENTIAL DISTRIBUTION

A. Clifford Cohen, University of Georgia, Athens,
Georgia;
Frederick Russell Helm, Georgia Southern College

1540-1720   TECHNICAL SESSION 4 - Room 372

Chairman: C. M. Greenland, Applied Mathematics
   Branch, Systems Analysis Office, Edgewood Arsenal
   Maryland

DYNOSS-DYNAMICALLY OPTIMIZED SMOOTHING SPAN

Roberto Fierro, White Sands Missile Range, New Mexico

1540-1720   TECHNICAL SESSION 4 (Cont'd)

                 STUDIES WITH A PROTOTYPE "OPTIMIZER" FOR USE IN
                 COMPUTER SIMULATION

                      Dennis E. Smith, HRB-Singer, Inc. Science Park,
                      State College, Pennsylvania

                 STAG MONOTONE EXPERIMENTAL DESIGN ALGORITHM (SMEDAL)

                      Private First Class Alexander Morgan, Systems Develop-
                      ment Division, US Army, STAG, Bethesda Maryland

1800-1900   SOCIAL HOUR - Officers' Open Mess

1900-   BANQUET - Officers' Open Mess

                 Presentation of the Samuel S. Wilks Memorial Award

                      Dr. Frank E. Grubbs, Chairman of the Conference
                      US Army Aberdeen Research and Development Center,
                      Aberdeen Proving Ground, Maryland

****

                      Thursday, 28 October

0830-1000   CLINICAL SESSION C - Sternberg Auditorium

                      Chairman: Captain Isaac S. Metts, Jr., Division of
                         Biometrics and Medical Information Processing,
                         Walter Reed Army Institute of Research, Walter
                         Reed Army Medical Center, Washington, D.C.

                 COMPARISON OF TREATMENTS GIVEN BI-VARIATE TIME RESPONSE
                 DATA

                      Pearl A. Van Natta, US Army Medical Research and Nu-
                      trition Laboratory, Fitzsimons General Hospital,
                      Denver, Colorado

                 EXPERIMENTAL DESIGN IN PROSPECTIVE STUDIES OF INFECTION
                 IN MAN

                      Major Robin T. Vollmer, US Army Medical Research
                      Institute of Infectious Diseases, Fort Detrick,
                      Frederick, Maryland

0830-1000   CLINICAL SESSION D - Room 358

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

0830-1000   CLINICAL SESSION D (Cont'd)

TREATMENT OF NULL RESPONSES

    Genevieve L. Meyer and R. L. Johnson, US Army
    Mobility Equipment, Research and Development Center,
    Fort Belvoir, Virginia

DESIGN OF RELIABILITY EXPERIMENTS TO YIELD MORE INFORMA-
TION OF FAILURE CAUSES

    Roland H. Rigdon, US Army Weapons Command, Artillery
    and Air Defense Weapons Systems Directorate, Weapons
    Laboratory, Rock Island, Illinois

0830-1000   TECHNICAL SESSION 5 - Room 372

    Chairman: Erwin Biser, US Army Electronics Command,
       Fort Monmouth, New Jersey

MAXIMUM LIKELIHOOD APPROXIMATION FOR GUMBEL'S LAW AND
APPLICATION TO UPPER AIR EXTREME VALUES

    Oskar M. Essenwanger, Physical Sciences Directorate,
    Research, Development, Engineering and MSL, US Army
    Missile Command, Redstone Arsenal, Alabama

STATISTICAL MODELS FOR HF IANOSPHERIC FORECASTING FOR
FIELD ARMY DISTANCES

    R.J. D'Accardi and R. A. Kulinyi, US Army Electronics
    Command, Fort Monmouth, New Jersey
    C. P. Tsokos, Virginia Polytechnic Institute and State
    University, Blacksburg, Virginia.

1000-1030   Break

1030-1220   CLINICAL SESSION E - Sternberg Auditorium

    Chairman: Henry A. Dihm, Jr., Aeroballistics
       LIRECTORATE, Research, Development, Engineering
       and Systems Laboratories, US Army Missile Command,
       Redstone Arsenal, Alabama

SOME PROBLEMS IN THE DESIGN OF TESTS TO CHARACTERIZE IR
BACKGROUND TRANSIENTS

    J. S. Dehne, J. R. Schwartz and A. J. Carillo, Combat
    Surveillance and Target Acquisition Lab, US Army
    Electronics Command, Fort Monmouth, New Jersey.

THE ANALYSIS OF A SUCCESS-FAILURE TIME SERIES WITH AN
APPRECIABLE NUMBER OF MISSING OBSERVATIONS

    Robert P. Lee, US Army Electronics Command, Atmospheric
    Sciences Laboratory, White Sands Missile Range, New Mexico

1320-1430	TECHNICAL SESSION 10 - Room 372 (Cont.)

Eugene Dupoit, Headquarters, US Army Munitions Command, Dover, New Jersey

1430-1500	Break

1500-1710	GENERAL SESSION 11 - Sternberg Auditorium

Chairman of General Session 11, Dr. Ivan R. Hershner, Jr., Office of the Chief of Research and Development, Washington, D. C.

CLASSIFICATION AND CLUSTERING TECHNIQUES IN DATA ANALYSIS

Dr. Geoffrey H. Ball, Senior Research Engineer, Stanford Research Institute, Menlo Park, California

HOTELLING'S WEIGHING DESIGNS

Professor K. S. Banerjee, University of Delaware, Newark, Delaware

****

Friday, 29 October

0830-1000	TECHNICAL SESSION 11 - Sternberg Auditorium

Chairman: Edmund H. Inselmann, US Army Materiel Command, Washington, D. C.

EXPERIMENTAL TESTING OF INTRUSION DETECTION DEVICES

Eric Mendelson, US Army Mobility Equipment Research and Development Center, Fort Belvoir, Virginia

LABORATORY CONTROL OF DYNAMIC VEHICLE TESTING

James W. Gran, US Army Tank-Automotive Command, Warren, Michigan

0830-1000	TECHNICAL SESSION 12 - Room 358

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

FUNCTIONAL PROPERTIES OF CSP-1 APPLIED TO A FINITE LENGTH PRODUCTION RUN

Richard M. Brugger, Concepts Branch, US Army Ammunition Procurement and Supply Agency, Joliet, Illinois

THE DEVELOPMENT OF FEATURE EXTRACTION TECHNIQUES FOR
PULSE DATA

    J. S. Dehne and Lieutenant B. E. Beaumont, Combat
    Surveillance, and Target Acquisition Lab, US Army
    Electronics Command, Fort Monmouth, New Jersey

0830-1000   TECHNICAL SESSION 13 - Room 372

    Chairman: Oskar M. Essenwanger, Physical Sciences
      Directorate, Research, Development, Engineering
      and MSL, US Army Missile Command, Redstone Arsenal,
      Alabama

THE USE OF A DESIGN OF EXPERIMENT FOR DETERMINING
OPTIMUM BRIGHTNESS OF PHOSPHORS EXPOSED TO HIGH-ENERGY
ELECTRON-BEAM BOMBARDMENT

    Joseph M. Velasquez and Isidore H. Stein,
    Quantum Electronics/Display Devices Technical Area,
    Electronics Technology and Devices Laboratory, US
    Army Electronics Command, Fort Monmouth, New Jersey

EXPERIMENTAL COMPARISON OF OPERATIONAL TECHNIQUES FOR
A SEMIAUTOMATIC FLIGHT OPERATIONS CENTER

    William Patterson, American Electronics Laboratories
    Company, Colmar, Pennsylvania
    Erwin Biser and Herman Menchar, US Army Electronics
    Command, Fort Monmouth, New Jersey

1000-1030   Break

1030-1220   GENERAL SESSION 111 - Sternberg Auditorium

    Chairman: Dr. Frank E. Grubbs, Aberdeen Research
    and Development Center, Aberdeen Proving Ground,
    Maryland

OPEN MEETING OF THE AMSC SUBCOMMITTEE ON PROBABILITY
AND STATISTICS

    Dr. Walter D. Foster, Analytical Sciences Directorate,
    Department of the Army, Fort Detrick, Frederick, Maryland

THE COMPARISON OF PROPORTIONS: A REVIEW OF SIGNIFICANCE
TESTS, CONFIDENCE INTERVALS AND ADJUSTMENTS OF STRATIFI-
CATION

    Dr. John J. Gart, Mathematical Statistics and Applied
    Mathematics Section, National Cancer Institute,
    Bethesda, Maryland

1220-1320   Lunch - *Officers'* Open Mess
****

RANDOMIZED RESPONSE: A NEW SURVEY TOOL
TO COLLECT DATA OF A PERSONAL NATURE

Bernard G. Greenberg
James R. Abernathy

Department of Biostatistics, University of North Carolina at Chapel Hill

Daniel G. Horvitz

Statistics Research Division, Research Triangle Institute,
Research Triangle Park, North Carolina

INTRODUCTION

Refusal to respond and an untruthful answer from a respondent are recognized
as two principal sources of nonsampling error that can influence sample estimates
involving surveys of human populations. These potential sources of bias are more
prevalent in surveys which involve sensitive or embarrassing questions. Warner[1]
the developer of the randomized response technique, designed the procedure in
order to reduce or eliminate bias introduced when respondents in surveys delib-
erately give false information or refuse to answer questions of a personal or
stigmatizing nature.

Randomized response is a relatively new and exciting statistical technique
of great potential in surveys of human populations. Although the method is still
in the infancy stage, much growth and development have occurred since its crea-
tion in 1965.

HISTORICAL DEVELOPMENTS

Qualitative response. In his original paper, Warner considered the case
where a proportion π of the population (say Group A) possessed some sensitive
characteristic while the remainder of the population did not possess the

---

-1-

characteristic. The objective was to estimate π. With the aid of a randomizing device, the respondent selects one of the following statements by chance,

I am a member of Group A.   [with a probability of P]

I am not a member of Group A.   [with a probability of (1-P)]

and answers "Yes" or "No" to whichever one of the two statements is selected. The interviewer does not know which statement was actually selected through the randomization process and, therefore, does not know to which statement the respondent's reply refers. Only the respondent knows to which one of the two statements his reply is addressed.

The rationale underlying the randomized response procedure is that, since the respondent can answer a sensitive question without revealing his personal situation, potential stigma and embarrassment on the part of the respondent are removed. Under these conditions there is no longer any need to refuse to respond or to give an incorrect or evasive answer. If respondents are convinced that the procedure guarantees anonymity in this respect, it follows that cooperation and validity of response should be improved.

The randomized response procedure clearly does not permit interpretation of *individual* responses and, in fact, is not designed for that purpose. Its sole purpose is to determine π in a *group* of respondents. For analytical purposes the total sample may be subdivided according to age, race, and other available characteristics in order to study the relationship of the various values of π in the subgroups.

Warner showed that the maximum likelihood estimate of π is unbiased, if persons are encouraged by the technique to tell the truth, and its value is

$$\hat{\pi} = \frac{1}{2P-1} \left\{ P-1 + \frac{m}{n} \right\}, \quad P \neq \frac{1}{2}, \tag{1}$$

where

n is the sample size of which m persons reply in the affirmative.

The variance of the estimate, as given by Warner, is

$$Var(\hat{\pi}) = \frac{\pi(1-\pi)}{n} + \frac{P(1-P)}{n(2P-1)^2}$$

Abul-Ela et al[2] extended Warner's model to the trichotomous case designed to estimate the proportions of three related, mutually exclusive groups, one or two of which possessed a sensitive characteristic. The model was further extended to estimate any j proportions (j>3) when all the j group characteristics are mutually exclusive, with at least one and at most j-1 of them sensitive. The reason behind this extension was to provide theory for the multichotomous situation which is often found in practice. The solution to this problem was found to lie in choosing a new, nonoverlapping sample with a different value of P for every additional parameter to be estimated.

Following a suggestion by Simmons, Abul-Ela[3] investigated and described a variation of the Warner technique known as the unrelated question model. As indicated above, the Warner technique concerned two questions (or statements) which are inversely related to the sensitive characteristic. The unrelated model is predicated on the assumption that confidence in the anonymity of the technique would be increased if two unrelated questions were used, one pertaining to the sensitive characteristic and the other to a nonsensitive, innocuous condition. Horvitz, et al[4] further discussed this modification and presented results from two field studies. Greenberg et al[5] studied the theoretical framework of the unrelated question model with respect to estimation of the sensitive attribute, variance of the estimates, effect of untruthful reporting, selection of the unrelated characteristic, and other design properties. Mean square error efficiency of the unrelated question model versus the original Warner model was reported to

favor the unrelated question model. Further reduction in mean square error was shown to accrue when the frequency of the unrelated neutral characteristic was known in advance. Furthermore, a method of formulating the unrelated question which insures *a priori* knowledge of its parameters was described. This latter suggestion was made to the senior author by Richard Morton of the University of Sheffield and it was field-tested by Greenberg et al[10] with favorable results. The results of this application are shown in the next section in an example where data on emotional problems are described.

The maximum likelihood estimate of the sensitive attribute, let us call it π_A now, in the unrelated question model (assuming that the proportion of the population possessing the unrelated characteristic, π_Y, is known in advance) and its variance are

$$(\pi_A | \pi_Y) = \frac{\lambda - \pi_Y(1-P)}{P}$$

$$\text{Var}(\hat{\pi}_A | \pi_Y) = \frac{\lambda(1-\lambda)}{nP^2}$$

(2)

where

λ = probability that a "Yes" answer will be reported, and

P = probability that the sensitive question is selected by the respondent.

To obtain an estimate of π_A, one simply substitutes the observed value of λ, say $\hat{\lambda}$, for its expected value in the above equation. $\hat{\lambda}$ is defined as $\frac{m}{n}$ where n and m have the same definition as in equation (1).

Although highly desirable, it is not mandatory that an estimate of π_Y be known in advance. There may be situations where π_Y is not known and not

amenable to *a priori* estimation. Under these conditions the sample design can be altered to permit estimation not only of the sensitive characteristic by π_Y as well. This may be accomplished by selecting two independent samples of sizes n_1 and n_2 from the population. There are two devices and they would be arranged in such a way that the probabilities of selecting the sensitive question in the two subsamples (P_1 and P_2, respectively) are different. The resulting proportions of "Yes" responses in the two subsamples, $\hat{\lambda}_1$ and $\hat{\lambda}_2$, would also be different. Estimation formulae, assuming π_Y not known in advance, are

$$\hat{\pi}_A = \frac{\hat{\lambda}_1(1-P_2) - \hat{\lambda}_2(1-P_1)}{P_1 - P_2}$$

$$\hat{\pi}_Y = \frac{\hat{\lambda}_1 P_2 - \hat{\lambda}_2 P_1}{P_2 - P_1}$$

with variances,

$$V(\hat{\pi}_A) = \frac{1}{(P_1 - P_2)^2} \left[\frac{\lambda_1(1-\lambda_1)(1-P_2)^2}{n_1} + \frac{\lambda_2(1-\lambda_2)(1-P_1)^2}{n_2} \right]$$

$$V(\hat{\pi}_Y) = \frac{1}{(P_2 - P_1)^2} \left[\frac{\lambda_1(1-\lambda_1)P_2^2}{n_1} + \frac{\lambda_2(1-\lambda_2)P_1^2}{n_2} \right]$$

Gould et al[6] considered the unrelated question method with two trials *per* respondent in an attempt to develop models of respondent behavior and apply these models to survey results on illegitimate births in which randomized response *was* used.

Barksdale[7] extended previous work in randomized response to permit the estimation of the coefficient of correlation of two related traits, at least one of which is stigmatizing. The potential value of estimating the proportion of individuals with two attributes was also treated.

Quantitative response. Most of the randomized response work to date has been concerned with refining the technique for use in personal interviews using questions of a qualitative nature requiring only a "Yes" or "No" response. The technique need not be restricted to nominal scale data, however. It has wide application in the area of quantitative response and study is being directed toward further development of the method in this area. Greenberg et al[8] have presented theory underlying the quantitative application of the randomized response procedure including unbiased estimates of the population mean and variance of both the sensitive and non-sensitive distributions.

The quantitative model differs from the qualitative in that the questions are designed to elicit a response in quantitative terms rather than "Yes" or "No". The responses to both the sensitive and non-sensitive question must be of the same relative magnitude, thus making it impossible to identify from the response which question has been answered. The overall distribution of responses is comprised of numerical answers to both questions. This distribution must be separated in order to provide estimates of the parameters of interest, viz., the population mean of both the sensitive and non-sensitive distribution, and their variances.

Assuming two independent, non-overlapping samples of sizes n_1 and n_2, unbiased estimators for the mean of the sensitive and non-sensitive distributions, μ_A and μ_Y respectively, are shown[8] to be

-6-

$$\hat{\mu}_A = \frac{(1-P_2)\bar{Z}_1 - (1-P_1)\bar{Z}_2}{P_1 - P_2}$$

(3)

$$\hat{\mu}_Y = \frac{P_2\bar{Z}_1 - P_2\bar{Z}_2}{P_2 - P_1}$$

where

P_1 = probability that the sensitive question is selected by the respondent in sample $i(i=1,2)$, $P_1 \neq P_2$

$\bar{Z}_i$ = mean response of sample $i(i=1,2)$

The variances of the two estimators, $\hat{\mu}_A$ and $\hat{\mu}_Y$ were found to be

$$V(\hat{\mu}_A) = \frac{1}{(P_1-P_2)^2} [(1-P_2)^2 V(\bar{Z}_1) + (1-P_1)^2 V(\bar{Z}_2)]$$

(4)

$$V(\hat{\mu}_Y) = \frac{1}{(P_2-P_1)^2} [P_2^2 V(\bar{Z}_1) + P_1^2 V(\bar{Z}_2)]$$

where

$$V(\bar{Z}_i) = \frac{1}{n_i} [\sigma_Y^2 + P_i(\sigma_A^2 - \sigma_Y^2) + P_i(1-P_i)(\mu_A - \mu_Y)^2].$$

The variances of the estimators can be conveniently estimated by using the sample variances s_i^2 in (4): $\hat{V}(\bar{Z}_1) = s_1^2/n_1$ and $\hat{V}(\bar{Z}_2) = s_2^2/n_2$.

Other developments. Since many surveys are self-administered by means of mail questionnaires or other media, extension of randomized response to such sources of information is essential. The North Carolina group of investigators is currently exploring the development of a randomization

device which can be implemented in the questionnaire. This technique presents no substantive problems in statistical theory different from those already studied but does present problems in question design, format, respondent acceptance, and illiteracy.

FIELD APPLICATIONS

The first field application of the unrelated question technique was conducted by the Research Triangle Institute in October, 1965,[4] and involved personal interviews in a total sample of 104 white and 44 black households in which it was known that a live birth had occurred within the previous two months. It was also known that 18.9 per cent of the births were illegitimate according to the birth certificate on file with the North Carolina State Board of Health. The respondents were asked to select a card from a shuffled deck of 50 cards and to answer "Yes" or "No" to the statement printed on the card. The two statements used in the deck were:

1. There was a baby born in this household after January 1, 1965, to an unmarried woman who was living here.

2. I was born in North Carolina.

The objective was to estimate the proportion of all households with a birth to an unmarried woman. The results were in remarkable agreement with the known proportions of illegitimacy in each race in the *sample*.

North Carolina Abortion Study. The North Carolina Abortion Study[9,10] was a major field survey in 1968 to test the practicability of the randomized response procedure in eliciting information of a still more sensitive nature from the general population. Information was sought on a variety of subjects including illegally induced abortion, oral contraceptive use, emotional problems, and income. Each of these areas of interest utilized an independent

-8-

sample of women 18 years old and over from the combined population of five metropolitan areas of the state.

In deriving estimates of induced abortion, the randomization device consisted of a small, transparent, sealed, plastic box with two questions printed on the top, easily legible to the respondent:

1. I was pregnant at some time during the past 12 months and had an abortion which ended the pregnancy.

2. I was born in the month of April.

The first printed statement had a small, red ball in front of it; the second statement had a blue ball in the same location. Inside the box there were 35 red balls and 15 blue balls. The respondent was asked to shake the box of balls thoroughly, and to tip the box allowing one of the freely moving balls to appear in a "window" located in the device and clearly visible to the respondent. The color of the ball which appeared in the window determined which of the two statements the respondent answered with a simple "Yes" or "No". If a red ball appeared, she answered the abortion statement; if a blue ball appeared, she answered the born-in-April statement. Each respondent was urged to experiment with the plastic box several times to assure herself that both colors of beads could appear in the window. Color blindness is no problem with female respondents.

On the direction of the interviewer, the respondent then shook the box to answer the question officially. The interviewer was some distance away from the respondent so that she was unaware which statement had been selected. The respondent's reply of "Yes" or "No" was recorded by the interviewer with no knowledge concerning which statement had been answered.

Subsequent estimates of induced abortion in the population of women 18-44 years of age indicated that 3.4 per cent of them had experienced an induced

abortion during the previous year. By racial groups the estimates were 1.4 per cent for whites and 6.8 per cent for nonwhites. There is no way these findings can be absolutely validated. They do appear to be within the realm of reason, however, based upon limited evidence available for comparison.

The women who were asked about abortion during the past year were also queried about use of the contraceptive pill. The procedure used was exactly the same as that used for abortion, except that the statements were different, as follows:

   1. I am now taking "the pill" to prevent pregnancy.

   2. I was born in the month of April.

The results indicated that one-fourth of the women were taking oral contraceptives at the time of the survey. A higher estimate was found for nonwhite women (28.4%) than for white women (23.1%) and this concurred with previous expectations because of family planning programs conducted in some areas for poverty groups.

In another sample designed to obtain information on emotional problems, the following set of questions was used with women 31 years old or over:

   1. At some time during my life I had an emotional problem which caused me to seek help from a professional person, such as a psychiatrist, doctor, clergyman, psychologist, or social worker.

   2. The color of the ball in the window is blue.

The emotional problem trial was different from the others in that it used three colors of balls: red, white, and blue. After shaking the box in the usual fashion and causing a ball to roll into the window, respondents with a red ball in the window answered the emotional problem statement. Those with either a white or blue ball in the window answered the other statement.

Estimates of the proportion of women with an emotional problem showed that 22.8 per cent had experienced such a problem which required professional attention. The estimates were higher for nonwhites (30.6%) than for whites (19.9%). As with abortion and oral contraceptives, the emotional problem estimates were computed under the assumption that the value of the unrelated question for the group was known in advance. This value for the emotional problem aspect of the study was defined as the ratio of the number of women drawing a blue ball to the number drawing a non-red ball. The exact value of the numerator and denominator were, of course, unknown. The expected value of this proportion is the ratio of blue balls to non-red balls in the randomizing device, and this was known. It was considered to be a reasonable estimate of the proportion answering "Yes" to the statement, "The color of the ball in the window is blue" and was used in calculations of emotional problem estimates.

The North Carolina Abortion Study also collected *quantitative* data on abortion and income using the randomized response procedure. The statements on abortion were as follows:

1. How many abortions have you had during your lifetime?

2. If a woman has to work full-time to make a living, how many children do you think she should have?

The principal difference between this application of the technique and the situations previously described is that the answer was expected to be a small number rather than a "Yes" or "No" response. The estimation procedure was entirely different, of course. The income statements were completely analogous to those for number of abortion; in both application and analysis. They were:

-11-

1. About how much money in dollars did the head of this
   household earn last year?

2. About how much money in dollars do you think the average
   head of a household of your size earns in a year?

Estimates and variances of the mean number of abortions obtained over a
lifetime in an urban population of women, and of mean income of heads of
households, were reported in [8]. A more detailed treatment of data collected
in the North Carolina Abortion Study and the analysis, both qualitative and
quantitative, may be found in [9,10].

Other applications. Proctor and Gamble, Ltd., conducted a survey in
England[11] in 1968 using the randomized response technique to gain informa-
tion on personal hygiene habits of housewives. Two sets of questions were
used with the plastic box technique previously described:

Box 1

1. Have you cleaned your teeth today?

2. Were you born in one of the following months:
   January, February, March, April, May, June, or July?

Box 2

1. Have you had a bath in the last two days?

2. Were you born on one of the following days of the month:
   1st, 2nd, 3rd, 4th, 5th, 6th, 7th, 8th, 9th, 10th, 11th,
   12th?

For purposes of comparison, each of the two "sensitive" questions above
was also asked through the direct question approach of an analogous group of
women. The two sets of estimates of the proportion of housewives cleaning
their teeth agreed within limits of sampling error. The same was true of the
estimates pertaining to having a bath in the past two days. The authors con-

-12-

cluded that the housewives understood what was required of them in the randomized response application, but that neither "sensitive" question proved to be embarrassing to the housewives. Hence, they felt the randomized response was unnecessary as long as direct questioning is not embarrassing to the respondent.

Barksdale reported using the randomized response procedure among selected students at the University of North Carolina[7] to justify the feasibility of some of the extensions to the method which he recommended as a result of his research. He was concerned with determining the proportion of students who had smoked marijuana at some time in their lives, and the proportion who had never cheated on an examination. This survey was beset with many unforeseen problems leading the author to conclude that "due to the lack of an adequate sample size and experimental procedure, little credibility can be associated with the observations".

The Research Triangle Institute used the technique in 1970 in a drinking-driving attitude survey in Mecklenburg County, North Carolina[12]. Comparisons were made between randomized response and direct question responses.

LESSONS LEARNED FROM FIELD EXPERIENCES

The randomized response technique has now been widely tested in the field and much has been learned as a result of these valuable experiences. Some of the additional knowledge gleaned from these surveys is of a positive nature which might be implemented immediately to improve estimates of the sensitive characteristic. The field experiences also identified problems in the procedure which are not readily amenable to solution, and which require further research. Some of the more important lessons gained from field experiences are enumerated below.

1. The respondent must be convinced that the survey is for legitimate
    research purposes, and that his anonymity to the sensitive question
    cannot and will not be violated. The institution with which the
    interviewer identifies himself is of particular importance in that
    the respondent must not get the impression that some agency is prying
    into his personal life for some ulterior motive.

2. Estimates of the parameter, π_A, and its variance are extremely sensitive
    to the P_1 (and P_2) values (probabilities of selection of the sensitive
    question associated with the randomization device). It is, therefore,
    imperative that the device *in practice* produce these values within reasonable
    sampling error and untainted by non-sampling error and bias.

3. Further research should be directed toward determining optimum P_1
    (and P_2) values from the point of view of acceptability to the respondent.

4. Estimates of the parameter, π_A, and its variance are sensitive to the
    choice of the alternate (or non-sensitive) question. The latter should
    be chosen with known and relatively low prevalence in the population
    being surveyed. It should not be foreign to the group, and must be culturally
    acceptable. It is best to choose π_Y in the anticipated region of π_A. If these
    conditions cannot be met, consideration should be given to formulating
    the alternate question as a function of the randomization device (as
    was done in the emotional problem trial).

5. The interviewer should thoroughly understand the technique and be able
    clearly and lucidly to impart this knowledge to a respondent whose
    education is minimal. The interviewer must evoke a feeling such that
    the respondent has confidence in her integrity and sincerity. Furthermore,
    the interviewer should have faith in the procedure and not be skeptical
    lest this attitude be unconsciously transmitted to the respondent.

1. The respondent must be convinced that the survey is for legitimate research purposes, and that his anonymity to the sensitive question cannot and will not be violated. The institution with which the interviewer identifies himself is of particular importance in that the respondent must not get the impression that some agency is prying into his personal life for some ulterior motive.

2. Estimates of the parameter, π_A, and its variance are extremely sensitive to the P_1 (and P_2) values (probabilities of selection of the sensitive question associated with the randomization device). It is, therefore, imperative that the device *in practice* produce these values within reasonable sampling error and untainted by non-sampling error and bias.

3. Further research should be directed toward determining optimum P_1 (and P_2) values from the point of view of acceptability to the respondent.

4. Estimates of the parameter, π_A, and its variance are sensitive to the choice of the alternate (or non-sensitive) question. The latter should be chosen with known and relatively low prevalence in the population being surveyed. It should not be foreign to the group, and must be culturally acceptable. It is best to choose π_Y in the anticipated region of π_A. If these conditions cannot be met, consideration should be given to formulating the alternate question as a function of the randomization device (as was done in the emotional problem trial).

5. The interviewer should thoroughly understand the technique and be able clearly and lucidly to impart this knowledge to a respondent whose education is minimal. The interviewer must evoke a feeling such that the respondent has confidence in her integrity and sincerity. Furthermore, the interviewer should have faith in the procedure and not be skeptical lest this attitude be unconsciously transmitted to the respondent.

6. The problem of administering the technique to illiterates must be
   resolved.

7. The technique should not be used in situations where the direct question
   can be used.

8. Research toward a more efficient randomization device should continue.
   Making P equivalent to $\frac{1}{2}$ has intuitive appeal and would tend to lessen
   suspicion on the part of respondents that the procedure was rigged in
   some way. This value cannot be used in the original Warner model as
   indicated in Equation (1); it can be used in the unrelated question
   model (Equation 2) but with a loss in efficiency when compared with
   higher values of P. This loss might possibly be overcome by increased
   cooperation on the part of respondents when, for example, a simple toss
   of a coin is used in lieu of the box-and-balls device currently used.

POSSIBLE APPLICATIONS IN MILITARY RESEARCH

There are several sensitive areas involving the military in which informa-
tion may be vital but unavailable because, for obvious reasons, the direct
question approach would not yield truthful responses.

One such matter involves the use of drugs by members of the armed forces.
The randomized response method could be utilized to determine the proportion of
returning personnel or veterans, for example, who experimented with heroin,
or any of the other drugs available during their service abroad. Such a
study would permit estimates of drug use by members of the military classified
by rank, branch of service, location of station, age, length of service, race,
and a multitude of other characteristics.

Homosexual practices among the military is another field in which reliable
information may be practically non-existent. It is generally conceded that

knowledge of the extent of a problem is a prerequisite to launching a program to prevent or reduce the problem. This would be particularly true for homosexual or other deviant sexual practices that may exist in the armed forces. Detailed classification by branch, rank, and other strata would also be available.

A study of attitudes and opinions of members of the armed forces toward officers, military service in general, discipline, and related matters is another fertile field for exploration with the randomized response technique.

Alcoholism among the military is yet another field in which the lack of valid data on prevalence could be rectified through the use of randomized response. Obviously, there are many other areas of potential use in the military which could be enumerated but those mentioned suffice to indicate the potential of the method.

SUMMARY

A method of reducing or eliminating bias introduced when respondents in surveys deliberately give false information, or refuse to answer questions of a sensitive nature, is described. The method is known as randomized response.

The growth and development of the method since its origin in 1965 to the present time is documented. Special emphasis is placed on lessons that have been learned through several field experiences involving application of the technique. Some further research on the methodology is indicated.

Several potential applications of randomized response among the armed forces are discussed. These include the use of drugs, homosexual practices, attitudes or opinions toward military matters, and alcoholism.

REFERENCES

[1] Warner, Stanley L., "Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias," *Journal of the American Statistical Association*, 60 (1965), 63-9.

[2] Abul-Ela, Abdel-Latif A., Greenberg, Bernard G. and Horvitz, Daniel G., "A Multi-Proportions Randomized Response Model," *Journal of the American Statistical Association*, 62 (September 1967), 990-1008.

[3] Abul-Ela, Abdel-Latif A., "Randomized Response Models for Sample Surveys on Human Population," Unpublished Ph.D. Thesis, University of North Carolina, Chapel Hill, 1966.

[4] Horvitz, Daniel G., Shah, B. V. and Simmons, Walt R., "The Unrelated Question Randomized Response Model," *Proceedings of Social Statistics Section*, Washington, D. C.: American Statistical Association, 1967.

[5] Greenberg, Bernard G., Abul-Ela, Abdel-Latif A., Simmons, Walt R. and Horvitz, Daniel G., "The Unrelated Question Randomized Response Model: Theoretical Framework," *Journal of the American Statistical Association*, 64 (June 1969), 520-39.

[6] Gould, A. L., Shah, B. V., Abernathy, J. R., "Unrelated Question Randomized Response Technique with Two Trials per Respondent," *Proceedings of Social Statistics Section*, New York: American Statistical Association, 1969.

[7] Barksdale, W. B., "New Randomized Response Techniques for Control of Non-Sampling Errors in Surveys," Unpublished Ph.D. Thesis, University of North Carolina, Chapel Hill, 1971.

[8] Greenberg, Bernard G., Kuebler, Roy R., Jr., Abernathy, James R., Horvitz, Daniel G., "Application of the Randomized Response Technique in Obtaining Quantitative Data," *Journal of the American Statistical Association*, 66 (June 1971), 243-250.

[9] Abernathy, James R., Greenberg, Bernard G. and Horvitz, Daniel G., "Estimates of Induced Abortion in Urban North Carolina," *Demography*, 7 (February 1970), 19-29.

[10] Greenberg, Bernard G., Abernathy, James R., and Horvitz, Daniel G., "A New Survey Technique and Its Application in the Field of Public Health," *Milbank Memorial Fund Quarterly*, 48, Part 2 (October 1970), 39-55.

[11] Personal Communication to Senior Author from Mr. G. T. Park, Market Research Department, Proctor and Gamble, Ltd., Newcastle Upon Tyne, England, April 8, 1969.

[12] Gerstel, E. K., Moore, Paul, Folsom, R. E., and King, D. A., "Mecklenburg County Drinking-Driving Attitude Survey, 1970", Technical Report prepared under Contract No. FH-11-7538 for U. S. Department of Transportation by Research Triangle Institute, Research Triangle Park, North Carolina.

-17-

THE CHARACTERIZATION AND ANALYSIS OF
COMPLEX BIOLOGICAL RESPONSES

Captain J. Richard Jennings
Dept. of Experimental Psychophysiology
Walter Reed Army Institute of Research
Washington, D. C.

ABSTRACT. Complex, time-varying biological responses to discrete psychologica
events pose problems of numerical characterization and statistical analysis.
Biological responses, such as brief heart rate changes and cortical evoked
potentials, are multiphasic response with inherent dependency within the
response. This dependency should be identified in order to separate
biological control influences from the influence of the psychological
event. A prototype experiment is developed based on actual data illustrating
these problems. Alternate analyses are suggested and advice concerning the
application of multivariate techniques is sought.

INTRODUCTION. The accurate and complete numerical characterization of
a biological system's response to stimulation is a vexing problem for many
investigators. This is particularly true of the investigator seeking relations
between psychological and physiological functioning within a normal human
being. In such investigations a clearly specified psychological event is
presented and the physiological reaction is observed. The physiological
reaction is not simple, however. Reactions are generally extended in time and
vary during this time. In addition, such physiological reactions are multi-
determined -- biological controls on the response necessarily interact with
the reaction to the psychological event. Current knowledge of these response
determining events and their interactions doesn't seem to allow any straight-
forward deterministic modeling of the biological response process. The
first approximation to understanding these response processes must be, it
seems, through an empirical and statistical approach.

The biological responses of particular interest are those which occur as
time-varying responses to relatively specific stimulation. An example of
such a response would be a second to second change in heart rate induced
by a burst of white noise or an electrical potential change in brain waves
elicited by a flash of colored light. To elaborate the heart rate example,
at the onset of a burst of white noise, a person's heart rate, which
even without stimulation shows momentary variations in rate, may momentarily
slow down and then speed up prior to returning to a pre-stimulation rate.
With this type of response in mind, the basic questions of this paper can

Preceding page blank

-19-

be posed: "How can this biological response be characterized in a simple yet complete manner? How can the characterized response be compared to both pre-stimulus variability and responses to other psychological stimuli?"

A PROTOTYPE EXPERIMENT. A somewhat simplified prototype experiment has been designed that illustrates a general approach to experimentation with psychophysiology. After the description of this prototype experiment, the problem will be reduced to the comparison of two heart rate response curves based on real data collected under circumstances analogous to those of the prototype experiment.

The prototype study concerns the information processing of numbers within a map reading problem. Subjects highly familiar with a certain map were given two tasks. For both tasks a set of eight numbers representing map grid coordinates was presented visually for five seconds using a slide projector. In one task the subject was required to remember the exact coordinates. In the other task the subject was asked only to decide which quadrant of the map contained the coordinates displayed. The physiological concomitants of performance in these two tasks were of primary interest. More specifically, a pronounced physiological reaction was expected to accompany the task demand for exact grid specification, but not to accompany the less demanding quadrant placement task.

The experimental design specified the use of ten subjects. In order to use each subject as his own control, each subject performed in both tasks. Five subjects were randomly grouped to receive the coordinate memory task first and the remaining five received the quadrant placement task first. Three response measures were collected: heart rate, skin conductance (GSR's), and accuracy of performance on the information processing tasks. Each task was composed of ten trials using different map coordinates. For each task the response measures were averaged over the ten trials. The design can be diagrammed as follows:

COORDINATE MEMORY | QUADRANT PLACEMENT

Subject	Coordinate Memory RESPONSES			Quadrant Placement RESPONSES		
	Heart Rate	GSR	Accuracy	Heart Rate	GSR	Accuracy
Coordinate Memory Task First						
1	X_{1111}	X_{1112}	X_{1113}	X_{1121}	X_{1122}	X_{1123}
2	X_{1211}	X_{1212}	X_{1213}	X_{1221}	X_{1222}	X_{1223}
3						
4						
5	X_{1511}					X_{1523}
Quadrant Placement Task First						
6	X_{2611}	X_{2612}	X_{2613}	X_{2621}	X_{2622}	X_{2623}
7						
8						
9						
10	X_{2101j}					X_{21023}

Table 1. Idealized experimental lay out: X values represent averaged values over ten trials. Subscripting is represented by Xinjp, where i represents the 2 task orders, n represents the 10 subjects, j represents the 2 tasks, and p represents the 3 response variables.

Assuming for a moment that the three response measures were simply single numbers reflecting the heart rate, skin conductance, and performance responses, a rather straight-forward analysis of the results might be possible. A technique comparable to the profile multivariate analysis of variance suggested by Morrison (1967) might provide an overall estimate of the presence of any significant effect due to task and order treatments within the three response measures. Subsequent comparisons following a multivariate analysis could determine the locus of any significant effects. The application of this technique to the current data is not straight-forward. However, papers such as Potthoff & Roy (1964) suggest that generalization to cases such as the current one are possible. In short, if the data could truly be fit into the layout of Table 1, a comprehensive statistical analysis would seem possible.

The problem arises that, in fact, at least two of the response measures (heart rate and GSR) are not single numbers but a vector of dependent values reflecting both pre- and post-stimulus physiological activity. The decision of how to analyze this vector of values must be based on some knowledge of the nature of the physiological response and also upon the validity and availability of statistical methods of representing this data. A detailed knowledge of the heart rate response measure may aid in this decision.

THE HEART RATE RESPONSE. Psychophysiologists find it meaningful to measure heart rate during psychological performance because of a belief that the autonomic nervous system is continuously adjusting the internal mileau of the organism on the basis of events in the external environment. These adjustments are generally not massive enough to enter our awareness, although in states such as rage we are aware of physiological changes. Minor adjustments occur, however, and can consistently be related to psychological events. These adjustments appear as momentary changes that are not observed if pulse rate is measured over a minute or so. In order to observe such changes, the time between individual beats of the heart must be measured. This time can be converted to the familiar beats per minute (bpm).

$$\text{Heart Rate (bpm)} = \frac{60 \text{ seconds}}{\text{Inter-beat interval (seconds)}}$$

For example any interbeat interval of one second yields a heart rate of 60 beats per minute. In such a manner instantaneous heart rates can be expressed for individual beats before and after stimulation. In Figure 1 the rate of the heart is plotted for six beats prior to and six beats following stimulus presentation. Focusing on the curve with the solid line, the first three beats plotted show a stable heart rate of 82 beats per minute. This is followed by an increase to 83 bpm for the next three beats and then the reaction to the stimulus follows, a decrease in rate followed by a substantial increase. These are the sorts of brief changes in heart rate under study. Note that Figure 1 represents a total time of about ten seconds.

Returning to the prototype experiment, Figure 1 represents the cardiac responses to the two experimental tasks. The heart beat response curves are compared for the coordinate memory task and the quadrant placement task (the order of task presentation is ignored for sake of simplicity). The data

-22-
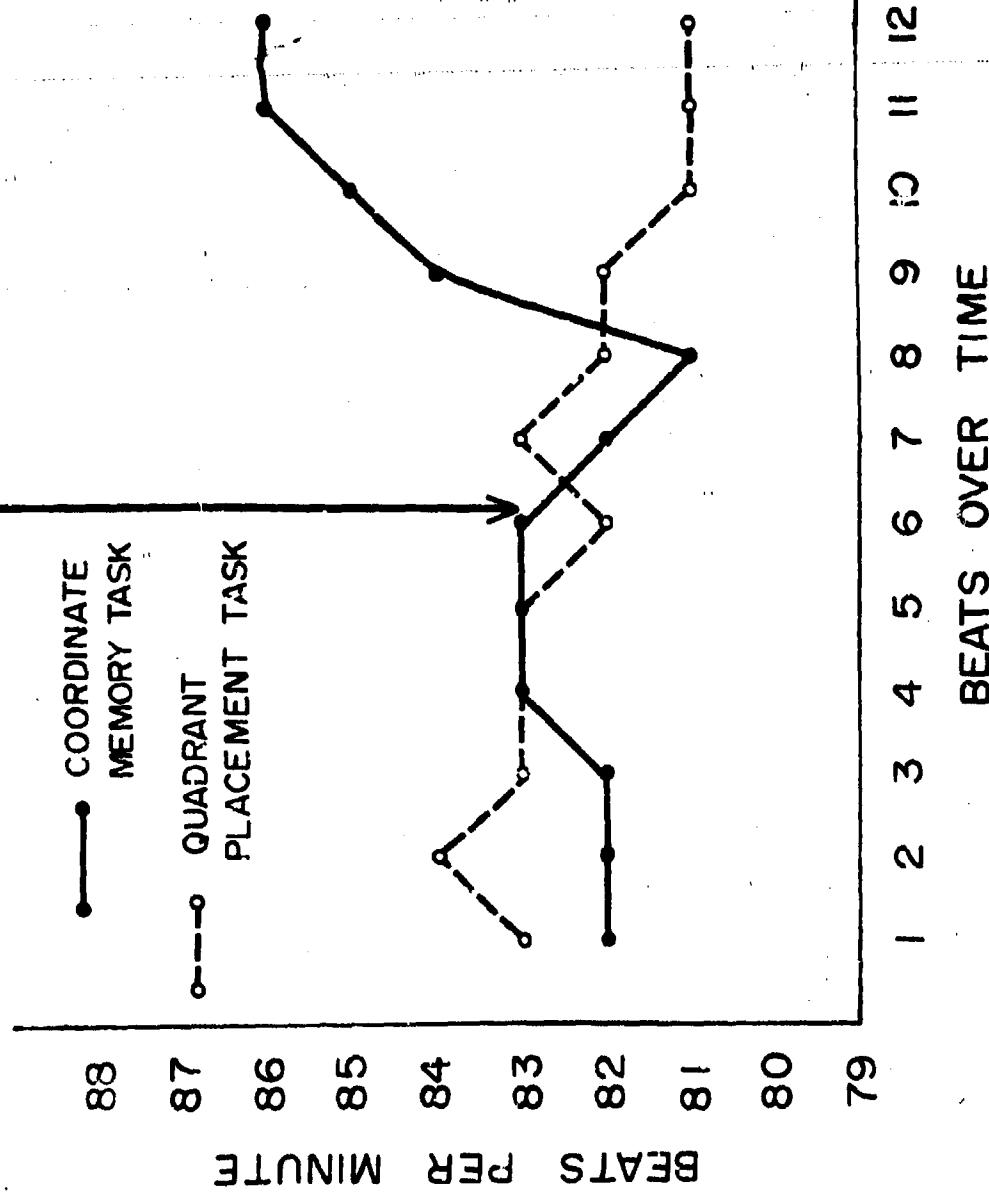
TABLE 2

Heart rate response data for the two information processing tasks*

SUBJECT	COORDINATE MEMORY												QUADRANT PLACEMENT											
	PRE-PERIOD						RESPONSE PERIOD						PRE-PERIOD						RESPONSE PERIOD					
	\(\bar{X}\) Value of Beat Number						\(\bar{X}\) Value of Beat Number						\(\bar{X}\) Value of Beat Number						\(\bar{X}\) Value of Beat Number					
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
1	78	78	78	79	80	82	74	76	81	83	83	82	82	83	81	81	80	77	80	79	77	74	73	73
2	94	95	95	94	94	92	90	89	92	93	94	94	92	92	93	94	92	91	91	92	94	93	94	93
3	79	80	81	84	95	85	78	78	82	82	87	83	79	79	79	79	81	82	79	77	77	78	78	83
4	8G	77	78	79	80	81	82	80	84	83	83	83	82	84	87	86	82	82	83	81	79	78	78	77
5	82	83	84	84	83	83	78	81	82	83	85	87	82	82	80	80	80	80	83	80	79	81	83	81
6	96	98	99	100	98	99	98	85	96	98	98	96	98	98	97	95	96	94	98	96	96	95	96	95
7	74	76	74	77	78	78	76	78	81	83	84	85	76	74	74	76	76	78	77	77	76	76	77	76
8	77	76	74	74	75	75	82	82	84	84	85	86	78	78	77	77	77	76	78	79	79	76	76	79
9	77	77	76	76	77	78	78	76	79	81	82	84	79	80	81	82	83	80	84	84	80	74	73	73
10	79	78	79	79	79	81	82	80	80	81	81	83	78	79	79	79	79	78	80	81	81	81	81	79
\(\bar{X}\)	82	82	82	83	83	83	82	81	84	85	86	86	83	84	83	83	83	82	83	82	82	81	81	81

*Values averaged over trials. Sessions effect is ignored for the sake of simplicity.

-23-

STIMULUS PRESENTATION

FIGURE 1

Averaged cardiac response curves for the coordinant memory and gradrant placement tasks.

-24-

for the individual subjects are shown in Table 2. Figure 1 indicates that during the coordinate memory task a relatively slow heart rate prior to stimulation was followed by a brief slowing and then a substantial speeding after stimulation. For the quadrant placement task a relatively fast heart rate was replaced by a progressively slowing heart rate after stimulation.

The basic questions of response characterization and analysis can be posed based upon the response curves shown in Figure 1:

1. Does the post stimulus reaction in either task constitute a real response to stimulation as opposed to a chance occurrence?

2. Given a true change in heart rate, do the complex, time varying responses in Figure 1 represent a unitary reaction of the cardiac control system? Alternately, can this reaction be meaningfully separated into a number of independent or semi-dependent components representing, perhaps, the influence of different cardiac control mechanisms?

3. Given a set of meaningful components, can a comparison across tasks be made of these components?

None of these questions are novel. Various approaches to their solution have been offered, but none have been generally accepted. Furthermore current interest in multivariate techniques leads to a reopening of such questions in the light of the availability of these methods. The eventual goal being a simultaneous analysis of a whole set of biological and psychological dependent variables.

SOME APPROACHES TO SOLUTIONS. Question 1: Does the post stimulus reaction in either task constitute a real response to stimulation as opposed to a chance occurrence?

A relatively straightforward solution to this problem may exist. The problem is essentially to decide whether two vectors both of size n (the baseline beats and the response beats) are significantly different. Given adequate experimental planning, an equal number of heart beats should be available before and after stimulation. The pre-stimulation values should be collected during a time period adjacent to the response when no psychological events are influencing the heart rate consistently. This is achieved by initiating stimulus events, e.g. map coordinates, at quasi-random times after the response to the previous event. Thus the stimulus induced perturbation in cardiac rate may be compared to a sample of normally varying cardiac rate.

Two requirements must be placed on the analysis. First, it must take into account the marked dependency between beats and between the two periods. In addition the integrity of the beats must be maintained -- averaging across beats would eliminate the transient response that is under study.

Two statistical approaches have been used both of which seem inelegant if not unsatisfactory. One approach (e.g. Coquery & Lacey, 1966) has been to do a series of univariate comparisons (e.g. Wilcoxon paired comparison or t-test for dependent x̄'s) between analogous beats (pre-stimulus beat 1 against post-stimulus beat 1, etc.). This approach yields a set of points in the response period that are supposedly significantly different from analogous points in the pre-period. This approach would seem to capitalize on the beat-to-beat dependency within the data as well as on random variations in the numerous errors terms used in the t-tests.

A second approach (e.g. Wilson, 1967 as derived from Greenhouse & Geisser, 1959) uses an analysis of variance design which treats beats as a classification factor. In the case of the current data, a two-way analysis of variance, might be set up with pre- vs. post-stimulation as a factor (two levels) and beats as factor (6 levels). A significant main effect of beats is interpreted as showing that consistent differences between beats occur. Central interest is, however, in the beats by pre- vs. post-interaction. This interaction is interpreted as showing that the beat-by-beat response curves are different between the pre- and post-stimulation period. Despite adjustments in degrees of freedom to "account for" dependency, this sort of analysis seems questionable if only because of the peculiar use of beats as a classification factor. Another problem with this analysis may be the varying degrees of dependency often found in physiological data.

The direction of solution for this problem seems at present to be a multivariate analysis comparing the two vectors simultaneously. The problem of comparing dependent vectors within the same sampling unit remains somewhat confusing, however. The intuitive appeal of the multivariate approach is the consideration of heart beats as a multiple response rather than a classification factor. The question remains whether the underlying model truly represents the data or questions better than the univariate approach. As noted previously, papers such as that by Potthoff and Roy (1964) suggest some optimism on this account.

Question 2: Given a true change in heart rate, do the complex, time varying responses in Figure 1 represent a unitary reaction of the cardiac control system? Alternately, can this reaction be meaningfully separated into a number of independent or semi-dependent components representing, perhaps, the influence of different cardiac control mechanisms?

If the cardiac response is indeed different from baseline activity, the nature of this response becomes the next question. Has the coordinate memory task produced a truly independent acceleratory change in heart rate or is the speeding seen in the latter part of the response period due to biological regulatory adjustment from the immediately prior slowing -- or even due to the relatively low pre-stimulation heart rate? In short what is the dependency structure within the total HR sample including baseline and response.

An example of the type of biological control system may be helpful in explaining the need to be wary of changes in response that are physiologically induced rather than caused by stimulation. This example will also demonstrate the need to look within the HR response for dependency. A basic and well known control mechanism present in biology is the homeostatic loop. In such cases differences from a preset internal state of the system are corrected by a negative feedback loop based on the output of the system. In the cardiac system, for example, baroreceptors in major arteries sense pressure and rate changes. Significant changes in rate or pressure cause the baroreceptor to send via the nervous system inhibitory signals which result in a corrective return of rate and pressure to normal levels. Thus, an efficient control loop prevents any major changes in cardiac parameters.

Within such a system, any change in heart rate would be expected to be followed by a feedback-induced change in the opposite direction. To the psychologist this feedback induced change is of little interest, but rather may mask to some degree further cardiac change induced by the experimental task. Thus a brief slowing of heart rate (as in the coordinate memory response curve) may elicit a homeostatic acceleration. This acceleration may either mask further deceleration of heart rate or combine with a psychologically-induced acceleration thereby confounding this reaction. Thus, if one wishes to demonstrate an accelatory reaction to a task, the independence of the acceleratory reaction vis-a-vis prior changes must be demonstrated. Prior changes must include both pre-stimulation oscillation as well as preceding changes within the cardiac response.

Psychologists have been concerned about biological control influences or reactions for at least 20 years. Early papers in the field formulated concepts such as the "Law of Initial Values" stating that the preresponse level of the variable influenced the response level. In general the notion was that the lower the initial level the greater the probability of a large response. In practice, most have followed the general outline of the suggestion of Lacey (1956) and empirically found the correlation between pre- and post-stimulation levels and covaried out this influence upon the response. Let us look at the application of such notions to our HR case.

The problem of the nature of the beat-by-beat data rises at once. The data is collected in order to look at momentary reactions not at changes from a base level to a response level. Through averaging across beats the HR data may be forced into this model, but there is no rationale for doing this. The underlying rationale for correcting for dependency is to subtract out or at least be aware of biological control influences upon the response unrelated to the experimental manipulations under study. The response assumedly due to the experimental manipulation is a momentary flunctuation and thus measures of this fluctuation not averages should be corrected. In like manner it seems more reasonable to covary with the pre-stimulus degree of fluctuation than to covary with the pre-stimulus average.

The influence upon one another of fluctuations within the response itself also should be investigated. Each task is expected, however, to produce a distinctive response curve. For example, the response to the coordinate

memory task includes an early deceleration while quadrant placement task does not show deceleration until late in the response period. The maximum heart rate in the coordinate memory task response should be corrected for the influence of prior deceleration. It is not appropriate, however, to perform the same correction on the quadrant placement data or indeed on the combined quadrant placement and coordinate memory data. In contrast the correction for pre-stimulation variability should be applied across all tasks. Pre-stimulus variability is assumed to be inherent variability and every pre-stimulus period should be a sample of this variability. Thus in order to correct for the influence of the inherent variability, the best estimate would be the mean of all the pre-stimulus periods.

The upshot of such thoughts seems to be an extensive correlational analysis of the complete HR (baseline and response) curve. Two analyses are suggested: 1) determine the correlation between pre- and post-stimulus fluctuation and then to remove the influence of the pre-stimulus fluctuation; 2) determine the structure of the post-stimulation response separately for each experimental treatment. The first analyses would involve correlations across subjects and across all experimental treatments while the second could be done within and/or across subjects within an experimental treatment group.

Both of the suggested analyses require that a score or set of scores representing HR change be derived from the data. These scores might be data points themselves or a derived statistic. Initially one might take points of intuitive interest such as maxima and minima and treat these as scores representing change. Two advantages of such scores can be cited. First, they estimate the fluctuation in the data without losing the identity of the maximum and minimum as would be done with a variance estimate. Second, they express the variability independently of the time dimension. This second advantage is important within the pre-period because this period is usually initiated randomly with respect to time. Thus the pre-period will show random segments of any regular periodicity. With respect to the post-stimulation response, the temporal information becomes important in defining the response, and thus simple maxima and minima are not as advantageous. In cases where the post-stimulation response is not precisely time-locked, however, the use of maxima and minima may be preferable to the method to be developed below.

The correlation of maximum and minimum scores from the pre and response periods could serve to estimate the degree to which the maxima and minima of the response reflect baseline fluctuations. This assessment may be made separately for the basal maxima and the basal minima's effect on response maxima and minima. Regression equations based on one or more of the four correlations computed could be used to "correct" the response scores by the baseline scores. This procedure appears to be a workable solution to the "initial value" problem albeit a cumbersome solution. Objections to this solution are the intuitive approach of looking at peaks and valleys and the probable unreliability of single points plucked from either the basal or response curves.

An alternative to the intuitive points of interest would be an attempt
to incorporate all the heart beats into the analysis. This would in general
mean doing analyses which preserve the time dimension. A simple correlation
between analogous points in pre- and post-stimulation response might be
considered. The random timing nature of the pre-stimulation period discussed
above argues against this approach for the pre- vs. post-response comparison.
Some form of time independent data representation, such as developed above,
would seem to be the preferred approach to the pre-post comparison. An
alternative might be to eliminate the specific pre-period in favor of a
substantial baseline period of ten minutes or so. Auto-correlational or
time series techniques could be applied to the baseline data to derive the
components of the resting variability. The response curves could then be
compared to the baseline spectra. This comparison would probably not be
statistical, however.

The dependency structure within the cardiac response to a specific task
(the second analysis) may be untangled with a time-based correlational analysis.
The fruitful application of this analysis requires that the cardiac response
to the task be time-locked in a similar manner across all subjects. If this
is the case, a simple correlation matrix consisting of each heart beat's
correlation with the remaining beats should be revealing. Beats, other than
immediately adjacent beats, that are significantly correlated would indicate
dependency within the response. Such dependency would be preliminary evidence
for either the operation of a biological control mechanism or the presence
of a unitary but multiphasic response. With such interpretations in mind,
it would be of interest to further explore the dependency structure. Factor
analysis might provide a statistical tool for this purpose. The ability to
isolate independent factors within the response is appealing conceptually
although practical problems of the interpretation of factors might emerge.

Other practical considerations are whether or not to rotate the factor
matrix, which rotation to use, and whether the factors should be derived from
the covariance matrix or the correlation matrix. (The covariance matrix is a
possibility because all of the scores have the same bpm metric). Even the
basic correlation matrix should answer, however, the question of whether
dependency exists with the cardiac response to a task.

In summary, an exploration of the question of the nature of the response
has suggested that this problem be split into two sections. 1) Can the
post-stimulation response be predicted from the pre-stimulation variability?
This question may be answerable by intercorrelating intuitive points of
interest, such as maxima and minima, across the two segment of the HR response
curve. 2) What are the dependency characteristics of the post-stimulation
response? The direction suggested to answer this question was an extensive
correlational analysis.

Question 3: Given a set of meaningful components, can a comparison
across tasks be made of these components?

The meaningful comparison of the cardiac response to quadrant placing and coordinate memory is the essential task remaining. Two of the previously suggested analysis techniques can be used to show some problems involved in such a comparison.

In discussing this question, the realistic assumption will be made that a moderate degree of dependency exists within the response to the task and between the pre-period and response period. If in fact little dependency exists, then relatively straight forward statistical comparisons are possible.

One approach to comparing cardiac responses would be to use the results of the factor analysis suggested earlier. Hopefully, this analysis produced two or three independent factors that were interpretable in terms of different influences on cardiac function. If one of these factors was interpretable as the influence of the psychological event, then primary interest would be in identifying this factor in the response to both tasks and then comparing the factor loadings. It is likely, however, that less clear factors would emerge and that more than one factor would seem implicated in the response to the psychological event. In either case, the factor structure of one task must be interpreted and then the factors within the response to the other task must be rotated to the criterion cf matching the structure of the first task response. Once this is achieved, a statistical comparison of the loading on each factor would be desirable.* It is very unclear whether some type of analysis of variance on factor loadings would be desirable and justified, (e.g., Factors X Task X Order X Replications). Indeed the whole procedure just outlined seems suspect and suggests that other forms of multivariate analysis might be preferable.

The use of intuitive peaks and valleys is another way of comparing task responses. These scores must be adjusted in some way by prior values, however. This necessity seems in general to lead to a proliferation of scores with dubious validities.

For example, a response maximum score may be corrected by covarying the pre-response maximum and minimum and, perhaps, even by covarying an immediately preceding response minimum. Seven different scores can be generated using different combinations of these covariates. These seven scores plus those for the response minima could be analyzed one-by-one using a univariate or multivariate analysis of variance. The high degree of intra response score dependency, however, poses interpretive problems. Similar argument can be made for and against a beat-for-beat comparison of two response curves.

---

* Dr. Robert Chapman also of the Department of Experimental Psychophysiology, WRAIR, has developed ideas along these lines in attempting to deal with the comparison of cortical evoked potentials. His ideas suggested this application to heart rate.

In conclusion, the problem of how to statistically analyze and characterize a complex, dependent time-varying response has been posed. The question was posed in terms of a prototype experiment and then a specific comparison of the responses specified the question in more detail. Tenative approaches to some aspects of the problem were offered but no truly coherent means of characterizing and analyzing the data could be offered. What would you do?

REFERENCES

1. Coquery, J. M., and Lacey, J. I. The effect of foreperiod duration on the components of the cardiac response during the *foreperiod* of a reaction time experiment. Paper presented at the meeting of the Society for Psychophysiological Research, Denver, October, 1966.

2. Greenhouse, S. W., and Geisser, S. On methods in the analysis of profile data. Psychometrika, 1959, 24, 95-112.

3. Lacey, J. I. The evaluation of autonomic responses: Toward a general solution. Annals of the New York Academy of Science, 67, 1956, 123-164.

4. Morrison, D. F. Multivariate Statistical Methods. New York, McGraw-Hill, 1967.

5. Potthoff, R. F., and Roy, S. N. A generalized multivariate analysis of variance model useful especially for growth curve problems. Biometrika, 1964, 51, 313-326.

6. Wilson, R. S. Analysis of autonomic reaction patterns. Psychophysiology, 1967, 4, 125-142.

COMBAT MODELS AS APPLIED TO RADIOTHERAPY*

Barry W. Brown
M. D. Anderson Hospital and Tumor Institute

James R. Thompson
Rice University and
R. M. Thrall and Associates

ABSTRACT

The equations of chemical kinetics have been applied to combat models for over half a century. Because of wide variability in military performance due to personal initiative, terrain, weather, etc., such applications have sometimes had their difficulties. However, kinetic equations are very useful in describing the treatment of malignant tumors by chemotherapy and radiotherapy. A fairly general model for radiotherapy is proposed. The combat analogy is that of an invulnerable blue side engaged in the destruction of a red force inextricably intermeshed with a defenseless pro-blue side, e.g., as in the use of air power to suppress a guerilla insurrection. It is demonstrated how the model answers some of the conjectures of experimental radiology. Moreover, the model shows how acute dose external beam treatment may be used to simulate continuous low dose therapy. This should enable (generally preferred where feasible) "radium" treatments to be employed in many more cases than at present.

<u>Introduction</u>. The credit for introducing differential equation models into the theory of combat is generally given to Lanchester [10]. *His* early two force models are of the general form

$$\frac{dx}{dt} = -a_1 x - a_2 y - a_3 xy$$

$$\frac{dy}{dt} = -b_1 x - b_2 y - b_3 xy \ .$$

These may easily be extended to a variety of heterogeneous force models.

Below we exhibit a combat model useful in formulating strategies in radiotherapy. In the context of warfare, it is a description of the use of airpower in a friendly area occupied by an enemy force. Let there be given an enemy with strength w and three friendly forces with strengths x, y, and z. Over a particular interval of time, let the equations of attrition be given by

$$\frac{dw}{dt} = -c_1 wz + q_w w$$

$$\frac{dx}{dt} = -c_2 xz + q_x x$$

$$\frac{dy}{dt} = -c_3 yz + q_y y$$

$$\frac{dz}{dt} = 0$$

where the replacement parameters satisfy $q_y > q_w > q_x \sim 0$.

Such a situation might arise where airpower, z, is to be used in destruction of an enemy with strength w . Because the enemy is mixed with the friendly inhabitants of the area, it is impossible to avoid a certain amount of non-enemy injury and loss of life. Some of this attrition will occur in segments of the

/

population - e.g., village elders, civil resistance forces, etc. - which can be satisfactorily replaced very slowly. These are the forces with strength x. Other non-enemy loss of life will include persons not in positions essential to the resistance effort. However ghastly their deaths are from a humanitarian standpoint, their niches in the society are filled relatively swiftly. These are the forces with strength y.. The enemy forces are considered to be replaceable rather slowly - e.g., by indoctrination of friendly forces. It is assumed that a fairly constant airpower committment is made by the friendly side with losses replaced immediately. The above model might be a useful description of a number of real world uses of airpower - e.g., some of the bombing campaigns in Indo - China. The problem is to minimize w subject to constraints on x and y .

It happens that the model described above is also useful in the consideration of the problem of obtaining optimal strategies in radiotherapy. The effect of a treatment on three separate cell groupings must be considered. First, there is the malignant tissue which we seek to destroy. Next, there are connective, vascular and nerve tissues which are replaced so slowly that they can be considered virtually irreparable. Finally, there are the functional tissues, skin, muscle, gut, etc. Under moderate radiotherapy, these tissues are replaced very swiftly and practically need not be considered as a constraint in the treatment. The problem, then, can be treated as one of constrained minimization. We wish to destroy as much neoplastic tissue as possible without causing so much destruction of connective, vascular and nerve tissues that a necrosis will result.

It is a generally accepted view that low rate continuous radiation by radium implants is preferred to high rate fractionated external beam therapy [4, 7, 12, 13]. Unfortunately, only a small percentage of malignant tumors can be treated by implantation.

The purpose of this paper is to use a Lanchester type model to give an interrelation between continuous and fractionated radiation therapies. Having achieved this objective, we shall demonstrate a scheme by which the effect of continuous therapy may be closely approximated by a fractionated regime.

Discussion. All our modelling has been carried out under the assumptions of multi-target theory (for an excellent introduction to the subject see Barendsen [1]). Basically, it is assumed that each cell has n targets. When all n targets have been damaged, the cell is destroyed. A convenient measure of dose is the number of D_0's delivered, where for a given type of cell a dose of one D_0 will result in as many hits as targets. All targets are repaired concurrently [5, 6, 20] and at a rate which is a parameter of the model. The equation for the surviving fraction of cells at the end of a high rate treatment of dose D (in D_0's) is given by

$$I = 1 - (1 - e^{-D})^n .$$

The change of sensitivity during the cycle [2, 8, 11, 15, 16, 17, 18] is accounted for by a variation in the target number [15]. This variation is approximated by breaking the cycle into two phases, one a resistant phase in which cells have a high target number, the other a sensitive phase in which a single hit kills the cell.

In Figure 1 we show our basic model. In a short interval of time, dt, a proportion $k_1 \, dt$ of the cells is leave resistant phase. During the same time interval, a proportion $k_2 \, dt$ leave the sensitive phase. During mitosis each cell is replaced by an average of q daughter cells which immediately enter the sensitive phase. When no cells die, q is 2.0. For nonproliferating tissue, q is 1.0. Generally, we have assumed a q value of 1.1 for neoplastic tissue.

It is assumed that during continuous radiotherapy, in addition to single target killing, there is an appreciable amount of single hit killing [1, 3, 9, 14, 19]. This damage may result from high LET traversal [3, 14, 19], the accumulation of a few relatively high energy traversals, or some target which cannot be repaired but which has a higher D_0 than the reparable ones. In systems with an appreciable number of targets, almost all killing from low rate radiation appears to be due to irreparable damage.

In continuous radiotherapy, killing is achieved at rates k_3 and k_4 from the sensitive and resistant phases respectively. Hence, k_3 is simply the dose rate (in units of D_c). Assuming that a fixed fraction ω of the dose delivered does irreparable damage, k_4 is ω times the dose rate.

To be more specific, let

$x^*(t)$ = the number of cells in the resistant phase at time t,

$y^*(t)$ = the number of cells in the sensitive phase at time t.

If we normalize these quantities (in order to talk about surviving proportions), we have

$$x(t) = x^*(t)/(x^*(0) + y^*(0))$$

$$y(t) = y^*(t)/(x^*(0) + y^*(0)) .$$

Then the kinetic equations for continuous radiation are given by

$$\frac{dx}{dt} = -(k_1 + k_4)x + k_2 qy ,$$

$$\frac{dy}{dt} = k_1 x - (k_2 + k_3)y .$$

The solution is given by

$$x(t) = a_1 \exp[(-c_1 + c_2)t] + a_2 \exp[-(c_1 + c_2)t] ,$$

$$y(t) = b_1 \exp[(-c_1 + c_2)t] + b_2 \exp[-(c_1 + c_2)t] ,$$

where

$$b_2 = (k_1 + k_4 - c_1 - c_2)/(-2qk_2c_2\frac{q}{T} - \frac{k_1 + k_4 - c_1 + c_2}{k_1 T})$$

T = cell cycle time

$$b_1 = \frac{1}{k_2 T} - b_2$$

$$a_1 = \frac{qk_2}{k_1 + k_4 - c_1 + c_2} b_1$$

$$a_2 = \frac{1}{k_1 T} - a_1$$

$$c_1 = \frac{k_1 + k_2 + k_3 + k_4}{2}$$

$$c_2 = \sqrt{c_1^2 - k_1 k_3 - k_1 k_2(1-q) - k_4(k_2 + k_3)}$$

At the end of the treatment $(t = t_e)$, the fraction of surviving cells is given by

$$f = x(t_e) + y(t_e) .$$

To obtain an algorithm for fractionated radiation, we use the above formulae during the time between treatments with $k_3 = k_4 = 0$. The surviving proportion at the end of a fraction is given by

$$x(t + 0) = x(t)[e^{-\omega D}(1 - (1 - e^{-(1-\omega)D})^n)]$$

$$y(t + 0) = y(t)e^{-D} .$$

-38-

Below, we alter the parameters in the model one at a time holding the *other* parameters at their "standard conditions" values (see Fig. 1). As previously noted, in a course of fractionated treatment the target number in the resistant phase is quite important. This fact can be seen from Figure 2. The surviving fraction increases dramatically as the number of targets increases from one to six. It then increases at a more leisurely rate to an asymptotic value where the only killing is of the single target variety in both the sensitive and resistant phases.

In Figures 3 and 4 we show the effect of the percent of the cell cycle spent in the sensitive phase on cell survival for both modes of radiation exposure. In both cases, the surviving fraction decreases with increasing time in the sensitive phase - more dramatically at low dose rates. In the fractionated course, higher dose rates always give lower surviving fractions. This is not so for continuous treatment. At some point, the effect reverses and low rates give a lower survival fraction than the high rates. The greater the proportion of the time in the sensitive phase the greater this effect, so that with forty percent in the sensitive phase the dose given at 5 D_0's per cycle produces a surviving fraction one-hundred times lower than when given at 25 D_0's per cycle. This might seem to be an argument in favor of dose rates lower than the 20 - 25 D_0's per cycle given in intracaviary therapy. It is doubtful, however, that the proportion of time in the sensitive phase approaches forty percent for any neoplastic tissues. Moreover, the points of crossover of the lines move to the right with increasing q . A gross lessening of dose rate could reduce the treatment effectiveness due to an increase in q during the period of treatment. Nevertheless, moderately lower dose rates than currently applied in radium therapy might very well be desirable in some cases.

The possibility of the production of lower survival fractions by low dose rates is, of course, due to redistribution. In the fractionated course, the

interval between treatments is half of a cycle, and redistribution of surviving cells into a mixture of sensitive and resistant phases is largely completed during this time. In a course of continuous radiation there is a different effect as shown in Figure 5. The fraction of cells in the sensitive phase falls rapidly early in treatment but quickly asymptotes. The value of the asymptote depends to a large extent on the rate at which cells in the compartment are killed. Hence, there is a greater fraction of cells in the sensitive phase at low dose rates. Consequently, less of the dose is wasted on cells in the resistant phase.

In Figures 6 and 7 we exhibit the effect of the proliferation rate, q, for continuous and fractionated strategies. A very important point which is observed is the fact that the efficacy of continuous treatment is little changed by even a drastic change in q. We conjecture that this may be an important factor in the generally preferred results obtained by radium therapy. In the fractionated course the surviving fraction changes by at least a factor of 100 as q goes from one to two.

In Figures 8 and 9 we examine the effect of the proportion of single hit events in the resistant phase. In the case of continuous radiation there is a great change in surviving fractions with ω. In the fractionated case, the change is much less.

In Figure 10 we see the effect of the number of high dose rate treatments per cell cycle on the surviving fraction of cells. The change in surviving fraction is large as we go from one to three treatments per cycle. However, for the dose rates considered there is essentially no change in surviving fraction with change in the number of courses per cycle if this number is at least seven. Beyond this number, only single target killing takes place.

Figures 11 and 12 give the surviving fraction as a function of the total dose delivered. There is little dose rate effect in the continuous case,

although the surviving fraction changes a bit more rapidly with total dose at
low dose rates. However, in the fractionated case, the dose rate effect is
significant, since multitarget killing increases with the dose rate.

In Figure 13 we note the lowering in the killing ability of fractionated
radiation if the target number in the resistant phase is increased to 10.

Let us now consider a practical example to interrelate continuous and
fractionated treatments. It shall be assumed that D_0 for both normal and
neoplastic tissues is 150 with a cell cycle time of 48 hours. It is further
assumed that $\omega = .3$, $q = 1.1$, the proportion of time in the sensitive phase is
.2, and the target number in the resistant phase is 6. The total dose of
radiation is limited by the allowable amount of killing of slowly cycling
tissues. A typical continuous treatment consists of a total dose of 7000 rads
over a period of 7 days by radium implants. This corresponds to 13.3 D_0's per
cycle and a total dose of 46.7 D_c's . The surviving fraction of the slowly
cycling connective, vascular and nerve tissues is given by

$$f = \exp[-\omega D_{total}] = e^{-14} \sim 10^{-6} .$$

The correponding surviving fraction of neoplastic tissue is obtained from
Figure 11 as slightly less than $10^{-7}(.44 \times 10^{-7})$.

Now let us consider a fractionated treatment of the same tumor with a
regime using an acute dose of 300 rads (4 D_0's) per cell cycle. We wish to
determine the total dose which would result in the same degree of destruction of
slowly cycling tissue as the radium stategy given above. This is easily obtained
by solving

$$(e^{-(.3)^2}[1 - (1 - e^{-(.7)^2})^6])^n = e^{-14} .$$

We obtain a solution of 17.4 treatments, or 34.8 D_0's (5200 rads). Using

Figure 12, we find a surviving fraction of malignant tissue slightly in excess of 10^{-7} (1.09×10^{-7}).

In this example the model gives neoplastic survival fractions for continuous and acute fractionated dose therapies which differ somewhat less than a factor of 3 in favor of the continuous strategies. It should be remembered that a constant value of $q = 1.1$ has been assumed throughout. Should q increase beyond this value during the course of treatment, then as Figures 3 and 4 demonstrate, the relative effectiveness of the continuous strategy will increase.

In general, the relative constancy of the killing ability of continuous radiotherapy with changes in the cell kinetics of the neoplastic tissue is a decided advantage over the usual fractionated strategies.

Inasmuch as radium implants can be employed in only a small fraction of tumors, it shall now be our task to demonstrate how the proposed model may be used to devise approximations to continuous treatment by fractionated external beam therapy. (The approximation of a constant valued function by a Dirac comb is, of course, a common technique in a number of fields, e.g., the design of filters in time series analysis.) For approximating the radium implant strategy above, it might be undesirable to deliver 1000 rads per day in small incremental acute doses, since this would involve, with present facilities, a constant wheeling of the patient in and out of the treatment room. (Perhaps in the future there will be a floor in major hospitals devoted to rooms for radiotherapy patients with cobalt machines, etc., located on tracks on the floor above so that it is possible to treat the patient in his own room.) However, delivering 1000 rads every two days in 125 rad doses every 6 hours should be feasible. This would be 6.65 D_0's per cell cycle instead of 13.30. An examination of Figure 11 shows that under the cell kinetic conditions listed, a

continuous strategy using 6.65 D_0's per cell cycle would actually increase some-
what the total killing of neoplastic tissue for a total dose of 46.7 D_0's (7000
rads). Figure 10 shows that at a dose of 6.65 D_0's per cell cycle, 8 equal
treatments per cell cycle give a slightly smaller surviving fraction than an
infinite number of treatments (i.e., continuous therapy) per cycle. Hence,
Figure 11 may be used to give conservative estimates of the surviving fraction
of malignant tissue. However, we must still determine the total dose allowable
under this regime to kill no more noncycling tissue than in the 7000 rad radium
strategy. Solving for m in

$$[\exp[-.3(125)/150] \, [1 - \exp[-.7(125/150)]]^8 \,]^m = e^{-14} \,,$$

we obtain a value of m = 56. Hence a total dose of 125 x 56 = 7000 rads is
allowable. Using Figure 11, we obtain a conservative value of $.44 \times 10^{-7}$ as the
surviving fraction of malignant tissue. Thus with the exception of better
control of dose distribution in the tumor by actual implantation, it would appear
that the effect of continuous treatment via radium implants may be nicely approxi-
mated by acute dose fractionated treatment.

Conclusions. We have seen that a Lanchester type combat model gives a straight-
forward means for interrelating continuous and fractionated strategies of
radiotherapy. At this stage in the modeling we have not taken account of the
dynamics of the model parameters during the course of treatment. Undoubtedly,
the length of the cell cycles, the proportion of time in the sensitive phase and
the value of .q change during the treatment. In addition, the pooling of all
cells with target numbers other than unity into a compartment with one target
number is an oversimplification. Moreover, the possibility of a growth fraction
of less than unity has not been considered. A mixture of cells, some stopped in
the resistant phase of the model and some cycling, can be handled. However,

pilot calculations indicated that if the growth fraction stays close to unity (say at least .9) the surviving fraction at treatment end is not much affected.

We wish to reiterate that calculations using models of solely multi-target killing by continuous low rate radiation predict much less destruction than is observed. In a fashion similar to others, we have postulated that a fraction of the dose delivered does single target killing regardless of the number of targets in the cell. It is not surprising that such a factor would have gone largely unnoticed, since it makes comparatively little difference in a fractionated acute dose regime. Contrary to one's initial reaction when first exploring the necessity for positive ω, it is not a bonus which makes continuous low rate radiotherapy feasible. One hit LET killing limits the flexibility of treatment, since without it one could much more fully exploit the differences between the kinetic parameters of neoplastic and normal tissues by adjusting dose rate and schedule.

The principle reason for lack of consideration of these matters is the fact that changes in cellular dynamics during treatment seem to be poorly understood. Moreover, it seems reasonable to study the simpler static parameter case before proceeding to the more complex dynamic parameter case.
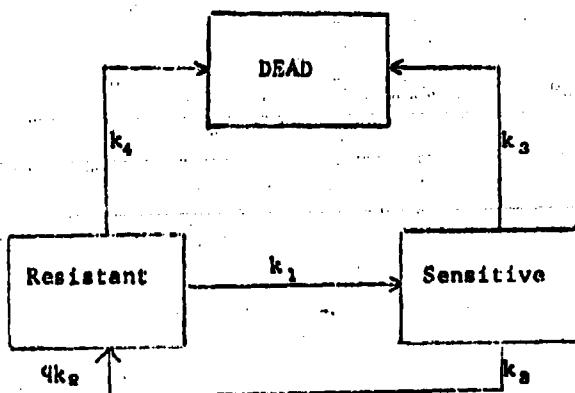
The model gives a plausible explanation for the fact that continuous radiotherapy has generally produced better results than fractionated acute dose treatment. Moreover, it gives a straightforward means for obtaining radium-like results by a regime of closely spaced fractionated treatments. Thus the advantages of radium treatment should be available to a much greater number of patients than is presently the case.

*Presently at the Massachusetts General Hospital.

APPENDIX

Figure 1

STANDARD CONDITIONS:

Treatment conditions -

 Course of treatment is 45 D_c's

 Proportion of radiation which does single hit killing is 0.3

 In fractionated high rate treatment, two fractions per cell cycle are

  given

Kinetic conditions -

 Proportion of time in sensitive phase is 0.2 of the cycle

 Proportion of time in resistant phase is 0.8 of the cycle

 The proliferation factor q is 1.1

 The target number is 6
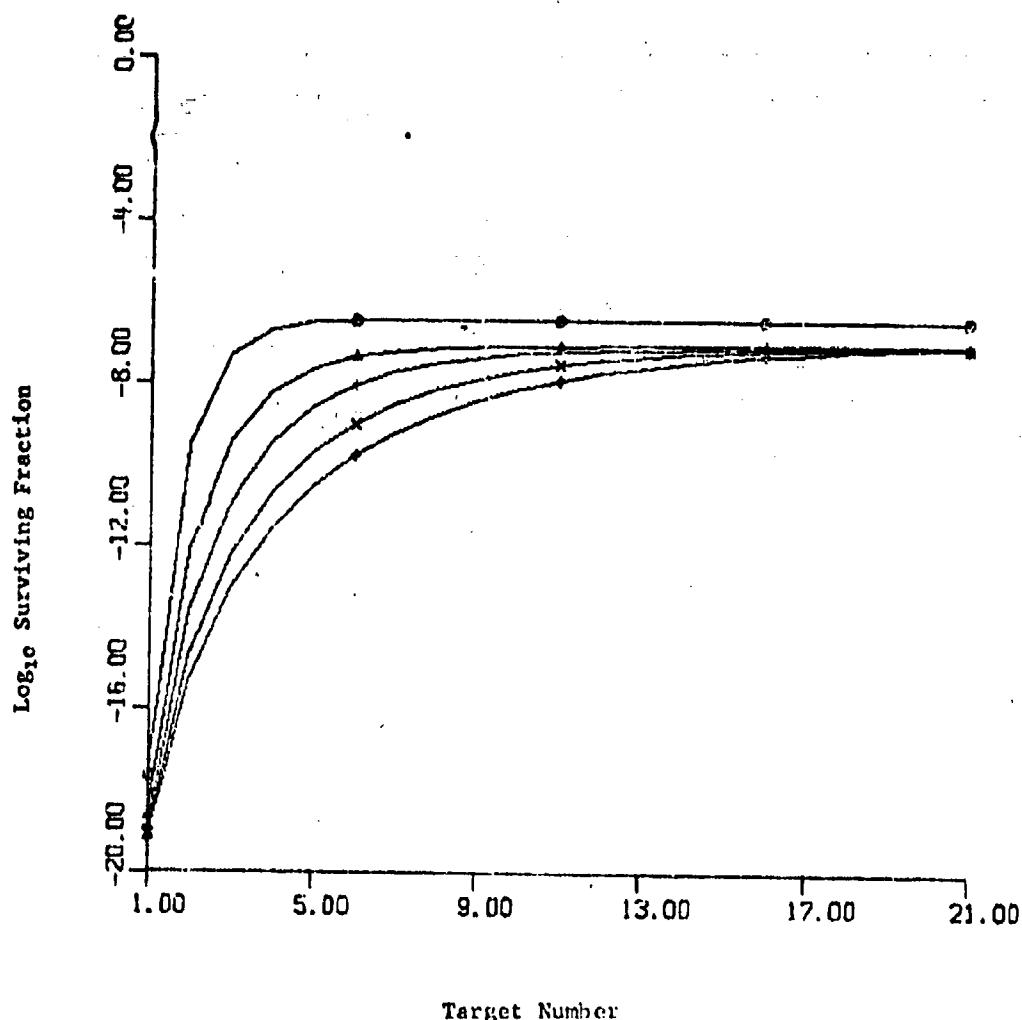
Figure 2

NUMBER OF TARGETS vs LOG SURVIVORS

(TWO FRACTIONS PER CYCLE)

$\omega = 0.3$

Percent of cycle in sensitive phase = 0.2

Total treatment dose = 45 D_0

$q = 1.1$

Dose rate top to bottom 1, 5, (1) D_0 per cycle.

Target Number

Figure 3

PERCENT OF CYCLE SENSITIVE vs LOG SURVIVORS

$\omega = 0.3$

Target number = 6

Total treatment dose = 45 D_0

$q = 1.1$
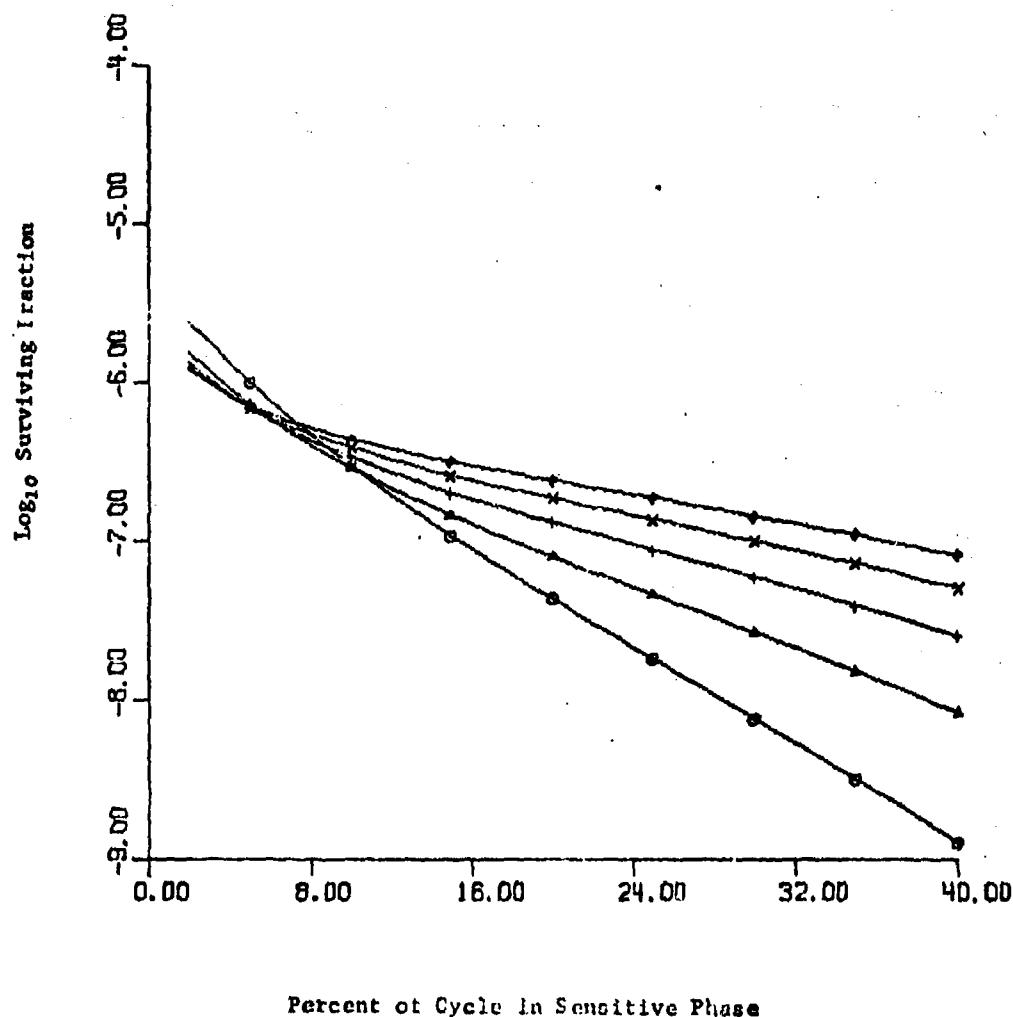
Dose rate top to bottom right 5, 25, (5)

Percent of Cycle in Sensitive Phase

-48-

Figure 4

PERCENT OF CYCLE SENSITIVE vs LOG SURVIVORS
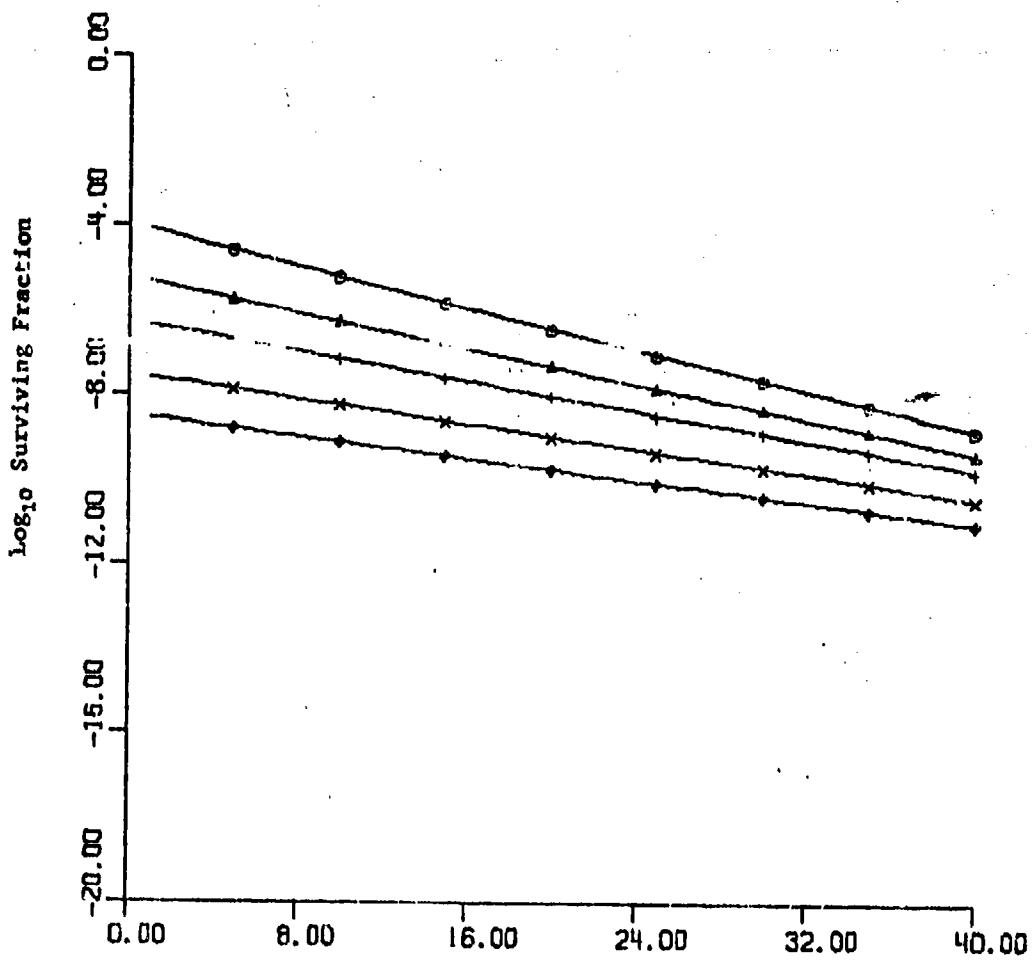
(TWO FRACTIONS PER CYCLE)

$\omega = 0.3$

Target number = 6

Total treatment dose = 45 D_0

$q = 1.1$

Doserate top to bottom 1, 5, (1)

Percent of Cycle in Sensitive Phase.

-49-

Figure 5

PROPORTION OF CELLS IN SENSITIVE PHASE vs DOSE

(CONTINUOUS)

$w = 0.3$

Target number = 6

Total treatment dose = 45 D_0

$q = 1.1$

Dose rate top to bottom 5, 25, (5)

Total Treatment Dose (in D_0 units)

-50-
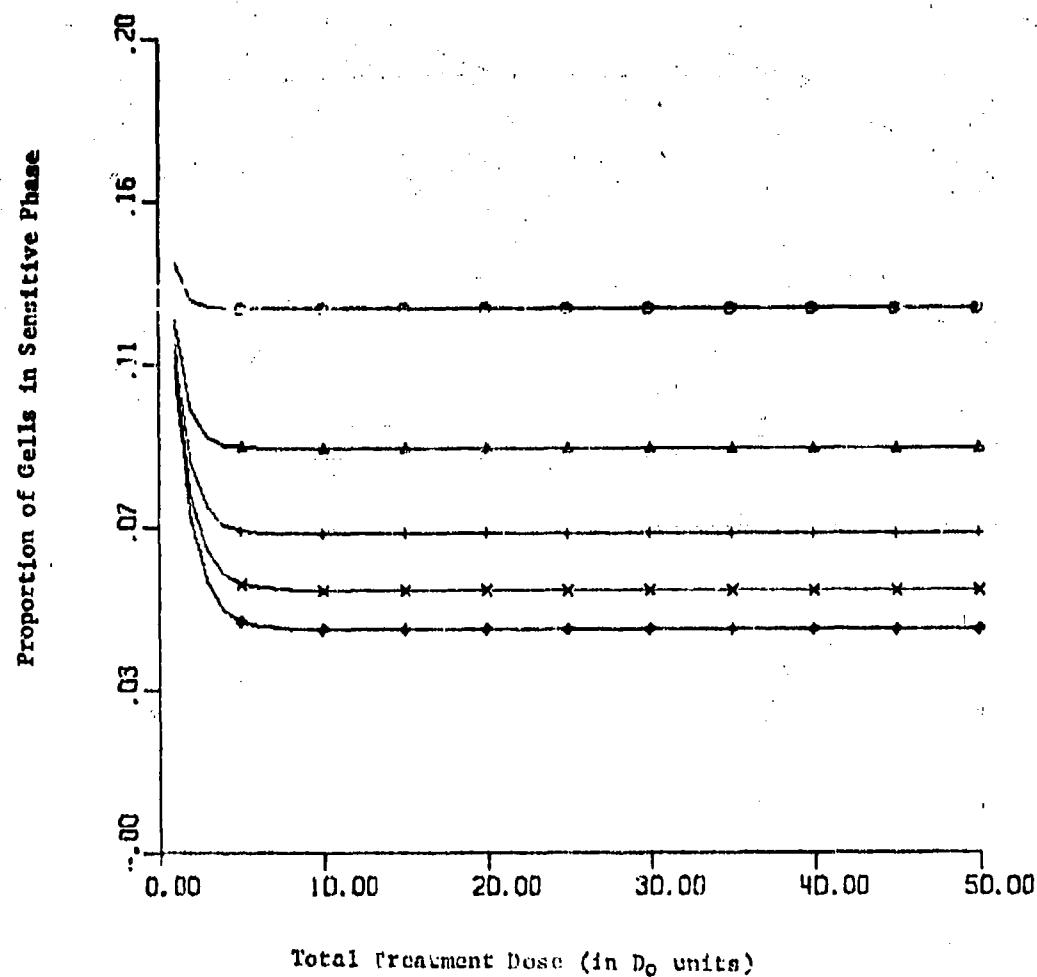
Figure 6

PROLIFERATION RATE vs LOG SURVIVORS (CONTINUOUS)

$\omega = .3$

Percent of cycle in sensitive phase = .2

Target number = 6

Total treatment dose = 45 D_0

Dose rate bottom to top left 5, 25, (5)
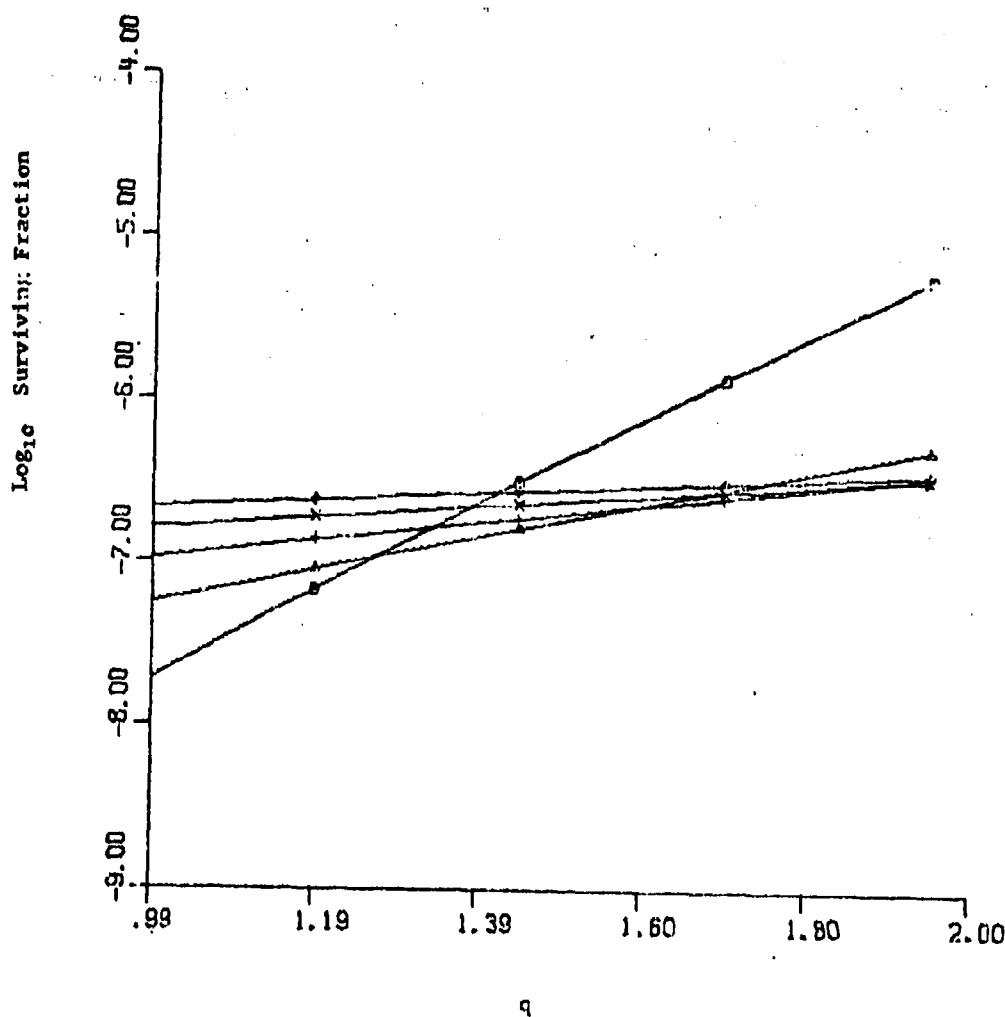
Figure 7

PROLIFERATION RATE vs LOG SURVIVORS

(TWO FRACTIONS PER CYCLE)

$\omega = .3$

Percent of cycle in sensitive phase = .2

Target number = 6

Total treatment dose = 45 D_0

Dose rate (top to bottom) 1, 5, (1)

q
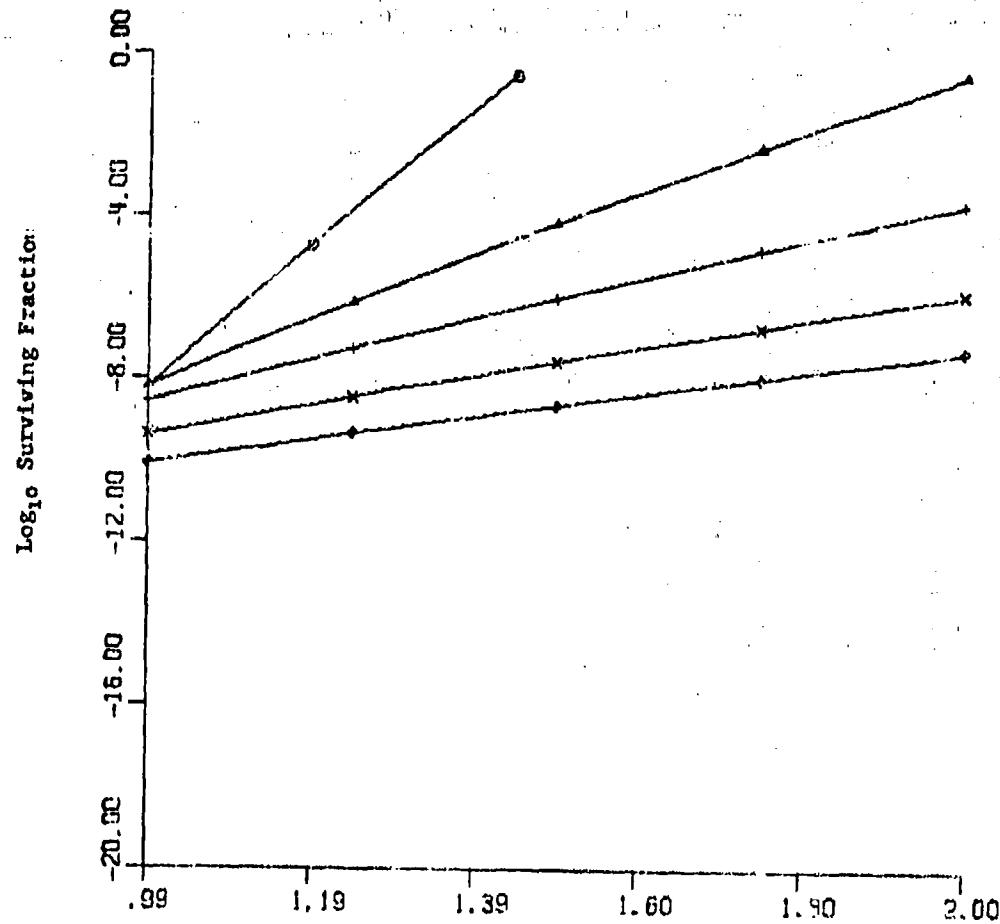
Figure 8

EFFECT OF ω ON TREATMENT (CONTINUOUS RADIATION)

Total treatment dose = 45 D_0

Proportion of cycle in sensitive phase = 0.2

Target number = 6
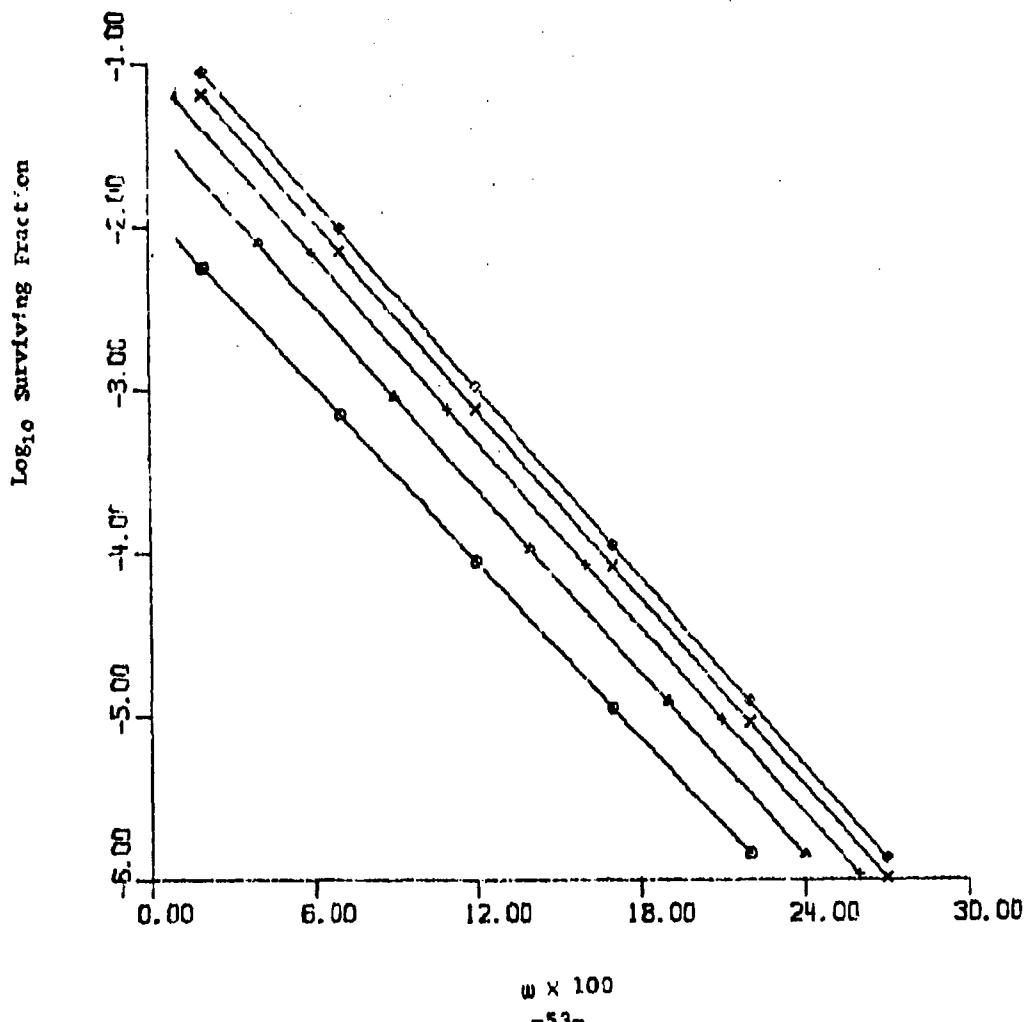
Dose rate from bottom to top 5, 10, 15, 20, 25, D_0/cycle

q = 1.1

Log$_{10}$ Surviving Fraction

$\omega \times 100$

Figure 9

ω vs LOG SURVIVING PROPORTION (FRACTIONATED)

Dose rate top to bottom 1, 5, (1)

Target number = 6

Proportion of cycle in sensitive phase = .2

Total treatment dose = 45 D_0

q = 1.1

ω x 100
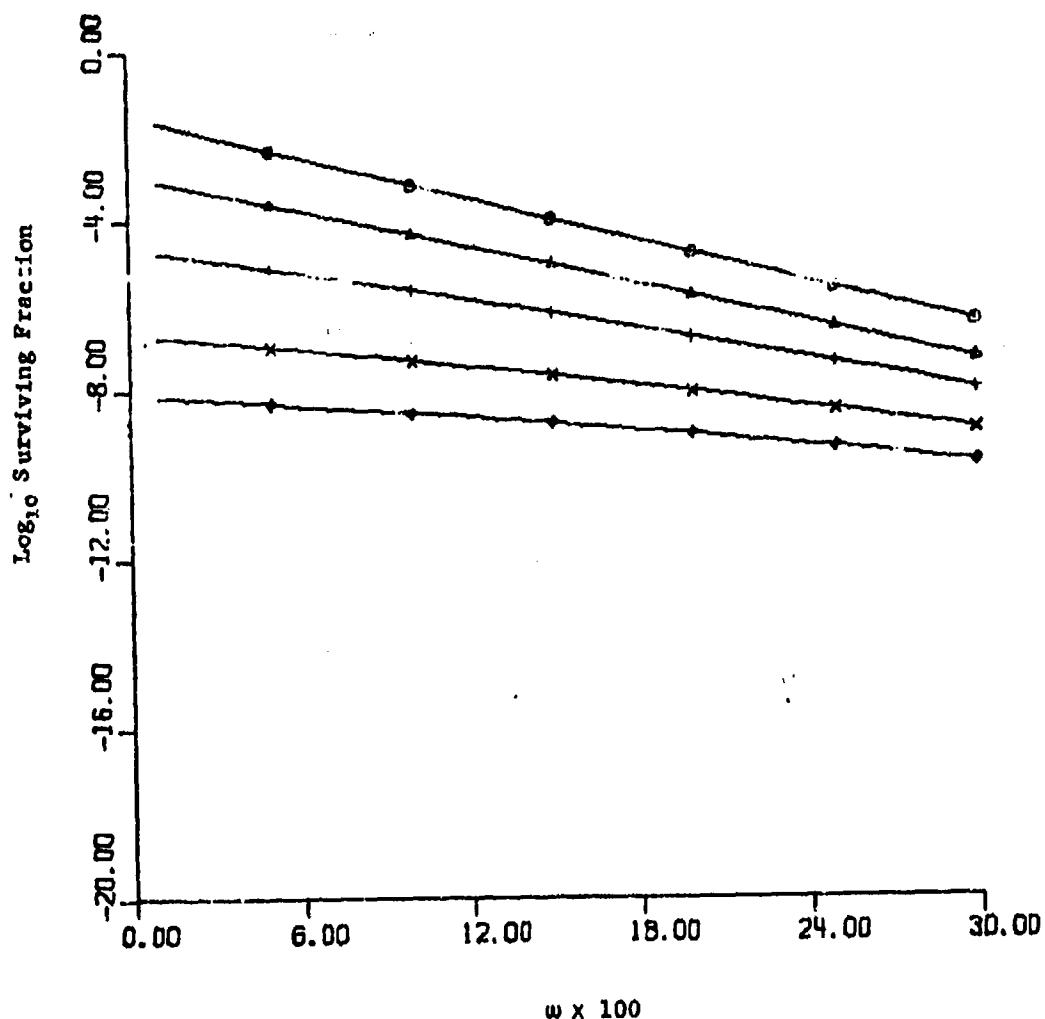
Figure 10

NUMBER OF FRACTIONS PER CYCLE vs LOG PROPORTION SURVIVING

$\omega = 0.3$

Target number = 6

Proportion of cycle in sensitive phase = .2

Total treatment dose = 45 D_O

q = 1.1

Dose rate top to bottom left 1, 9, (2)

Number of Fractions per Cycle
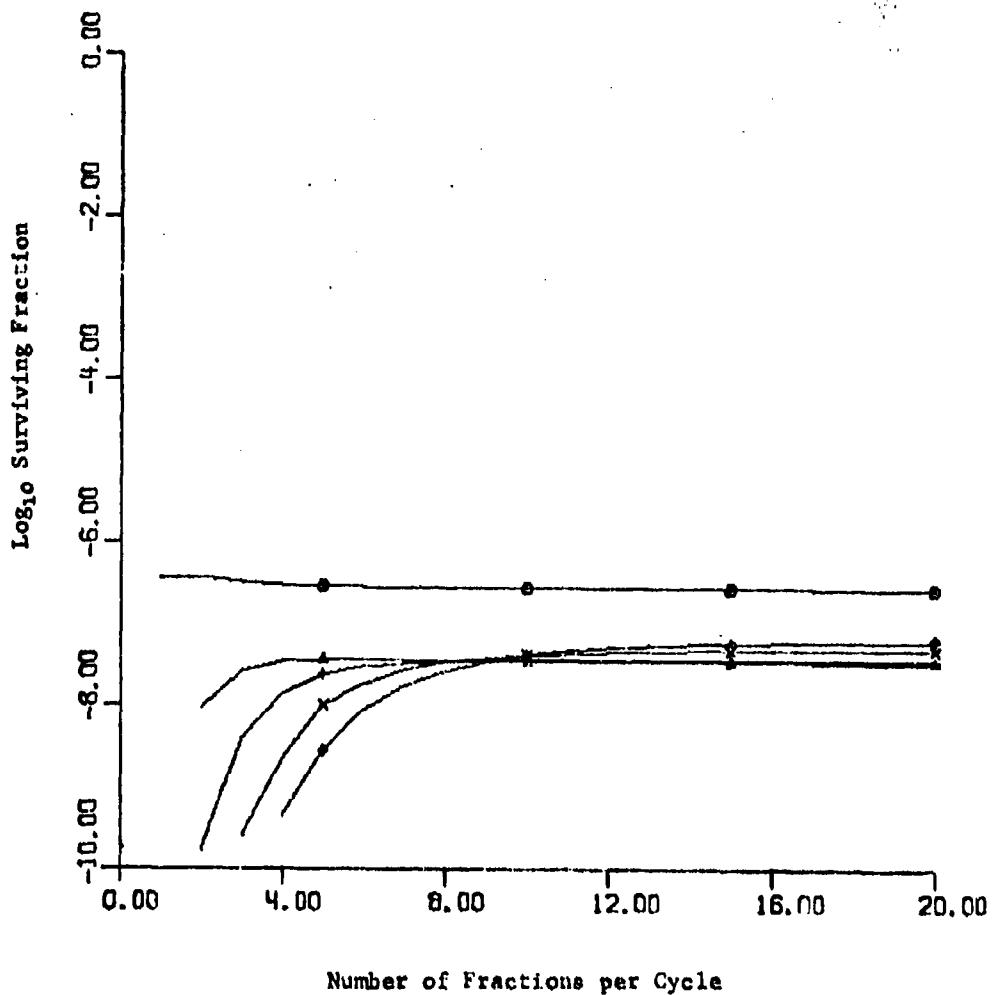
Figure 11

EFFECT OF TOTAL DOSE ON TREATMENT (CONTINUOUS)

$w = 0.3$

Target number = 6

Proportion of cycle in sensitive phase = .2

Dose rate 5, 25, (5)

Total Treatment Dose (in D_0 units)
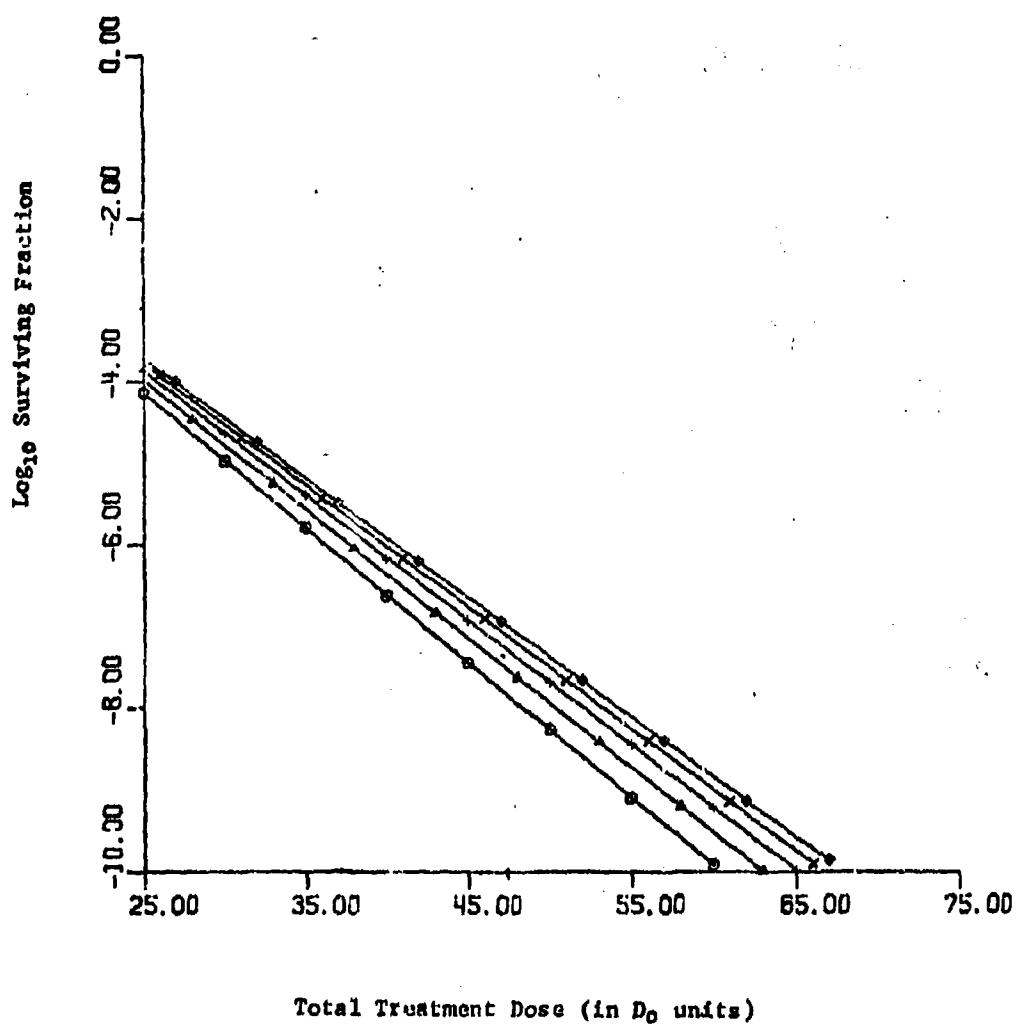
Figure 12

EFFECT OF TOTAL DOSE ON TREATMENT (FRACTIONATED)

TWO FRACTIONS PER CELL CYCLE

$\omega = .3$

Proportion of cycle in sensitive phase = .2
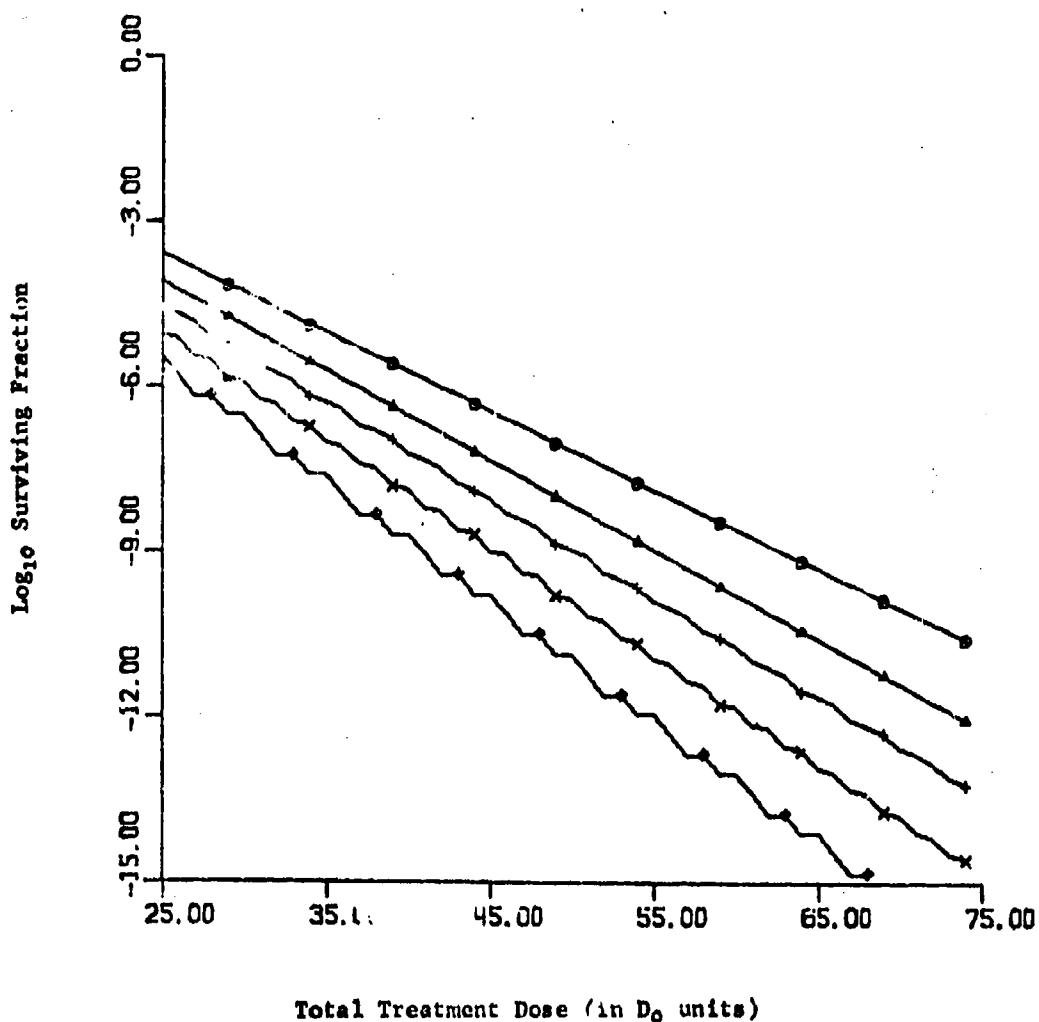
Target number = 6

Dose rate 1, 5, (1) D_0/cycle top to bottom

Total Treatment Dose (in D_0 units)

Figure 13

EFFECT OF TOTAL DOSE ON TREATMENT (FRACTIONATED)

TWO FRACTIONS PER CELL CYCLE
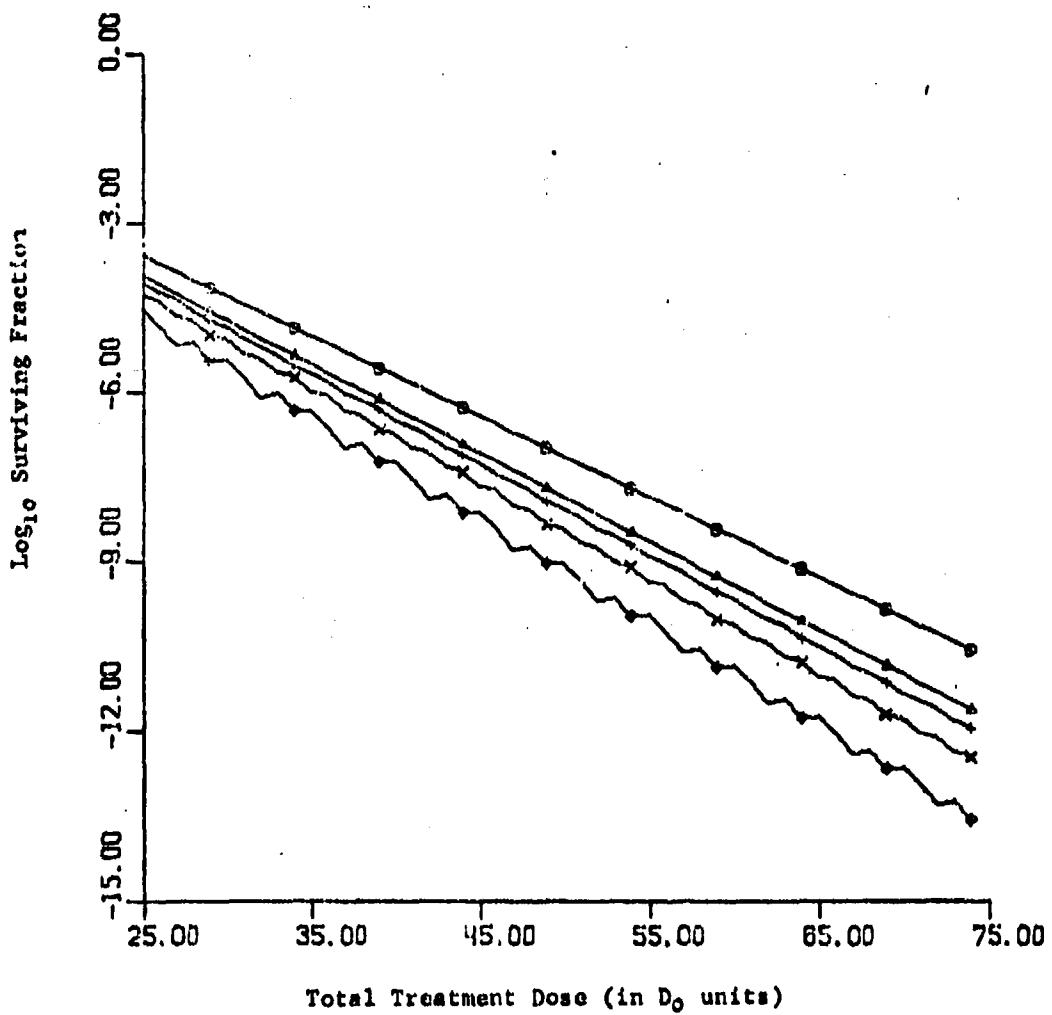
$\omega = 0.3$

Proportion of cycle in sensitive phase = .2

Target number = 10

$q = 1.1$

Dose rate 1, 5, (1) D_0 per cycle

REFERENCES

1. Barendsen, G. W., in Theoretical and Experimental Biophysics, pp. 167-23;, ed. Arthur Cole, Marcel Dekker, Inc., N.Y., 1967.

2. Burch P.R.J., "Some Aspects of Relative Biological Efficiency," Brit. J. Radiol., 30, pp 524-529, 1957.

3. Burch, P.R.J., "Calculations of Energy Dissipation Characteristics in Water for Various Radiations," Rad. Res., 6, pp. 289-301, 1957.

4. Cade, S., "Achievement of Radium in the Fight Against Cancer." Am. J. Roentgeol., 60, 713, 1948.

5. Elkind, M.M. and Sinclair, V.K., "Recovery in X-Irradiation Mammalian Cells, in Current Topics in Radiation Research, V. 1, ed. M. Ebert and A. Howard, pp. 165-220, North Holland Pub. Co., Amsterdam, 1965.

6. Elkind, N.M., "Sublethal X-ray Damage und its Repair, Mammalian Cells," in Current Topics in Radiation Research, ed. G. Silini, pp. 558-586, North Holland Pub. Co., Amsterdam, 1967.

7. Fletcher, G.H., in Textbook of Radiotherapy, pp. 110, 161-8, 179-180, 439 and others. Lea and Febiger, Philadelphia, 1966.

8. Gillette, E.L., Withers, H.R., and Tannock, I.F., "The Age-Sensitivity of Epithelial Cells of Mouse Small Intestine," Radiology, 96, pp. 639-663, 1970.

9. Hall, E.J., Bedford, J.S., "Dose Rate: Its Effect on the Survival of He La Cells Irradiated with Gamma Rays," Rad. Res., 22, pp. 305-315, 1964.

10. Lanchester, F. W., Aircraft in Warfare: the Dawn of the Fourth Era, Constable and Co., London, 1916.

11. Mauro, F. and Madoc-Jones, in Proc. Nat. Acad. Sci. U.S., 63, pp. 686-691, 1969.

12. Pierquin, B., Dutreix, A., "Toward a New System in Curie-therapy," Brit. J. Radiol., 40, pp. 184-186, 1967.

13. Regand, C., "Radium Therapy of Cancer at the Radium Institute of Paris, Technique, Biological Principles and Results," Am. J. Roentgeol. Rad. Therapy and Nuclear Med., 21, pp. 1-24, 1929.

14. Rossi, H. H., Biavati, M. H., Gross, W., "Local Energy Density in Irradiated Tissues, Radiobiological Significance," Rad. Res., 15, pp 431-439, 1961.

15. Sinclair, U.K., "Dependence of Radiosensitivity upon Cell Age in Time vs. Dose Relationships in Radiation Biology as Applied to Radiotherapy," Brookhaven National Laboratory, Pub. 50203 (C-57), pp. 97-107, 1970.

16. Terasima, J. and Tolmack, L.J., "Variations in Several Responses of He La Cells to X-Irradiation during the Division Cycle," Biophys. J., 3, pp. 11-33, 1963.

17. Thompson, L.H. and Humphrey, R.M., "Response of Mouse L-P 59 Cells to X-Irradiation in the G_2 Phase," Int. J. Radiation Biol., 15, pp. 181-184, 1969.

18. Whitmore, G.F., Gulyas, S. and Botond, J., "Radiation Sensitivity through the Cell Cycle and its Relationship to Recovery in Cellular Radiation Biology," Proc. 18th Am. Symp. Fundamental Cancer Research, Houston, March 1966, pp. 423-441, Williams and Wilkins, Baltimore, 1965.

19. Wideroe, R., "Radiobiological and Radiotherapy," Ann. N.Y. Acad. Sci., 161, pp. 357-367, 1969.

20. Withers, H.R., "Capacity for Repair in Cells of Normal and Malignant Tissues in Time vs. Dose Relationships in Radiation Biology as Applied to Radiotherapy," Brookhaven National Laboratory, publ. 50203 (C-57), pp. 54-69, 1970.

A COMPUTER PROGRAM FOR TRICHOTOMOUS BIOASSAY

Clifford J. Maloney
Fred S. Yamada
National Institutes of Health

A statistical technique is applied so frequently in a biological
context that it is widely described as bioassay, though it is by no means
confined to such applications and applications in other fields, e.g.,
explosive sensitivity or armor penetrability,[1] are common. The standard
text is that of Finney,[2] though here again the name, probit analysis,
applies only to one special approach and the book in fact discusses
other approaches. It is unfortunate that the terms "quantal regression"
or "quantal analysis," which are neutral, descriptive, and not preempted
for other purposes are not more widely adopted.

A second characteristic of the great majority of contributions to
bioassay (quantal regression) is that the response variable is assumed
to take just one of two values, response (which is often death or compo-
nent failure) and non-response. In some instances, however, the response
may take one of several forms which can be ordered as to severity. In the
case of an insect, for example, the response could be moribund as well
as dead. In the armor penetrability example, the response could be partial
as well as complete penetration. In such cases a method of analysis that
takes both types of response into consideration should give more precise
answers than the device that lumps intermediate responses either with the
responders or with the non-responders to permit application of standard
analysis.

A solution for this situation was supplied some years ago by
Burland et. al.,[3] giving explicit equations for the case of three classes
of response, the trichotomous case; though an appendix treats the general
situation of S outcome classes of one relation. The fully worked out

---

[1] Golub, Abraham, and Grubbs, Frank. "Analysis of Sensitivity Experiments
   When the Levels of Stimulus Cannot be Controlled."
   JASA, June 1956, Vol. 51, pp. 257-265.

[2] Finney, D. J. PROBIT ANALYSIS, Third Edition, Cambridge University, Press, 1971.

[3] Gurland, John, et. al. "Polychotomous Quantal Response in Biological Assay,"
   Biometrics, Vol. 16, No. 3, September 1960, pp. 382-398.

computing formulas for relative potency--including data for both a test product and a standard, each with two ordered classes of outcomes--are applied to an illustrative example. The particular example happens to have the same number of subjects exposed at all doses on one relation and for both relations--test and standard--though this is not necessary.[4] It may be helpful to some readers of the Gurland Paper to be told that doses given in Table 1 are to be converted to common logarithms, though the paper itself makes no mention of this step.

This paper reports the fact that the appropriate calculations have been programmed in the BASIC and FORTRAN languages and applied to the example in Gurland's paper. General availability of both programs is contemplated. These first programs are limited to the specific one cycle calculation described in (3). Modification to permit iterative fitting is modest. Finney applied his generalized computer program to this case also,[5] but his program is not directly applicable to currently available computers.

---

[4] Gurland, John. Personal Communication.

[5] ibid, pages 223-226.

CRITERIA FOR A BIOCELLULAR MODEL

George I. Lavin
Vulnerability Laboratory
BRL, ARDC
Aberdeen Proving Ground, Maryland

We are concerned with the design of a biocellular model. This model is to be used to evaluate the modification of performance of an animal system consequent to the absorption of energy by the animal.

Previously reported work has had to do with the application of specific, non-destructive analytical procedures for the assessment of traumatic consequences. (ARO-D Report 69-2)

A considerable effort has been expended on animal models. However, little attention has been given to the physiological aspects of the problem. Most of the publications have had to do with a mathematical type of analysis. This, I think, will be of value when we have learned more about the biochemistry, biophysics and bionics of the systems involved.

It is suggested that the criteria include:

The relationship between the biochemical composition and the specific functionality of the system(s) under consideration: Proteins, Nucleic Acids, Nucleo-proteins, Polysaccharides, Lipoids, Lecithins, Cephalins...

The mechanism of energy absorption by the system: The origin of spectra.

The equipartation of the absorbed energy: Formation of wound tracts. The initiation of atom and free radical reaction chains.

The dimensional-physiological nature of the systems: Soft tissue, Hard tissue, Heart, Kidney, Brain, Lungs, Muscle, Nerve, Circulatory systems....

Feedback effects: Secondard traumas. Shock.

SUMMARY

An attempt is being made to obtain a perspective of the reality of animal-task performance in terms which are involved in the performance.

FACTORIAL EXPERIMENTS OF SMALL ARMS WEAPON FIRE CONTROL

Adolph P. Kawalec
Quality Assurance Directorate, Frankford Arsenal,
Philadelphia, Pennsylvania

Frankford Arsenal, in Philadelphia, is engaged in the research and development of fire control instruments for a wide variety of uses such as sights, rangefinders and other instruments for helicopters, howitzers, tanks, artillery, mortars, recoilless rifles, as well as vehicle diagnostic equipment and computers. This paper deals with the test and analysis of small arms weapons fired from moving tracked vehicles.

Future track vehicle personnel carriers are being designed with ports to enable troops inside the vehicle to fire hand held weapons. Instability of the weapon caused by vehicle motion can be separated into two categories:

1. Lateral shifts in line of sight to other parallel lines of sight.

2. Change in direction of the line of sight.

Lateral vibration can change the point of impact on the target by only the distance over which the weapon moves. If the muzzle and butt are moved 10 millimeters in the same direction, impact will be 10 millimeters from where it would have impacted in the original position. If the muzzle moves 10 millimeters and the butt of the weapon remains stationary, the impact at a target at 100 meters would be 1 meter from the point at which the original line of sight would have caused it to impact.

This is the same change of angle characteristic of vibration which causes blur in cameras and optical viewing devices. Gyros, whose rotating mass directly resists motions which would cause a change in angle, have been used successfully in aerial photography for some time and have proven effective in reducing blur in military binoculars when used in moving track vehicles as recent tests performed by Frankford Arsenal have demonstrated.

A proposal was made by Frankford Arsenal to the US Army Small Arms Systems Agency in 1970 to perform a limited field evaluation of gyro stabilized weapons. The primary objective

---

The rest of this article was reproduced photographically from the manuscript submitted by the author.

Preceding page blank

was to determine the effect on hit probability of a weapon
with standard sights and a stabilizer when compared to the
same weapon without stabilization from moving vehicles.
Secondary objectives were to determine if a reflex sight
offered any improvement in hit probability with and without
stabilization from moving vehicles.

A Kenyan KS-6A stabilizer, which can be attached
externally to the weapon, was the stabilizer utilized
for this test. The Kenyan KS-6A was previously selected
from various other stabilization devices as having those
characteristics which make it ideal for use with small
arms weapons. Easily attachable, small size power supply,
etc. The stabilizer utilizes two encapsulated precision-
balanced gyros mounted in spring restrained gimbals which
resist pitch and yaw motion. The internal operation of
the gyro consists of heavy mass wheels which are rotated by
400 cycle, 155 volt AC motors at 20 to 22,000 RPM. The
stabilizer was attached to the pistol grip of the M16
Rifle. The reflex sight used was a commercially available
reflex sight with an electrically illuminated reticle dot.
The dot is projected to the plane of the target and reflected
to the eye by the front surface of the window through which
the target is viewed. The reflex sight was attached to the
carrying handle of the rifle and offers a simplification of
the sighting task by reducing the number of points to be
aligned for aiming from 4 to 2. Aiming is accomplished by
simply placing the dot of light on the target. The weapon
sight requires the alignment of the front sight, rear sight,
and eye with the target, whereas the reflex sight requires
that only the lighted dot reticle be aligned with the target.

TEST DESIGN

Normally, the type of test design is dictated by the
objectives of the test, the information sought and the cost
considerations associated with the degree of precision, sample
size and conduct of the test. The initial approach was to
consider a single factor approach to the testing of small
arms fire control. The single factor approach was rejected
since it was likely to provide a number of disconnected
pieces of information that would be difficult to piece
together to obtain conclusive results. For example, in order
to perform an experiment on a two-level factor (weapon sight,

-66-

reflex sight), some decisions must be made about the levels
of other factors: Speed, stabilization, shooters (shall
the fire control be studied at 8, 10, or 12 MPH in a moving
vehicle stabilized with shooter 1, 2, 3 or 4?). The
experiment would reveal the effects of the two-level factor
for this particular combination of speed, stabilization and
shooter.

It was not known whether or not the factors were
independent and since there were several levels of each
factor, a factorial approach was taken to examine the effects
on hit probability of the weapon sights and reflex sight at
various speeds, stabilization and with several shooters.
The factorial approach examines all the main factors and
their interactions simultaneously and would provide the
maximum amount of information to be returned. The tests of
hypotheses are:

    1. Each factor has no effect on the overall mean.

    2. There is no interaction between factors.

This factorial approach provides for the effects of the
two-level factor with any combination of speed, stabilization
and shooter that is included in the experiment and it was
expected that inferences could be made easily with this
design.

The test was divided into two phases:

    1. Stationary firing.

    2. Moving vehicle firing.

The test matrix for stationary firing measured the mean
radius of 10 round shot groups. This was a 4x3x2x2 factorial
planned to be replicated four times. Similar matrices were
also set up for the moving vehicle exercise. The moving
vehicle was separated into a "toward" target and "across"

target exercise. The data collected for the moving vehicle firing was the number of hits in each 10 round firing. This also was planned to be replicated four times.

Originally, each replication would be entered as a data point in the analysis, and since there would be equal subclass numbers (sample size), the computer program was written with this in mind before the completion of the test. After the testing was completed, several cases had more than four replications and in other cases less than four replications. The methods available of "filling in" missing data points was rejected as too laborious in terms of manual handling of the data and if computed would require a modification to the computer program. A major re-write of the program would also be required if the analysis was to consider the unequal replications. This dilemma forced a re-evaluation of the objectives. The 4x3x2x2 factorial would determine in the investigation of differences in means, differences of approximately 5% in hit probabilities. This type of difference is too small to justify the development of stabilizers or reflex sights for small arms use. Gross differences, on the order of 15% or 20%, would indicate that investment in stabilizers or reflex sight development could have a meaningful return (in terms of hit probability). Averaging the replications would not require a major computer program re-write and this method of "overcoming" unequal subclass numbers was finally selected with the full awareness of the loss in precision.

Figure 1 shows the moving target test layout. The across target firing zone was 100 meters from the target and 60 meters across. Five rounds were fixed in each direction by each shooter under each condition of the across target exercise.

The toward target course was 120 meters in length and varied in range from 170 meters to 50 meters from the target. Each shooter fired 10 rounds in the 120 meters of the toward target exercise. The vehicle speed was monitored by timing the vehicle between the "start to fire" and "cease fire" stakes. Precise vehicle speed was difficult to maintain and

controlled rather crudely with the vehicle　　　　ıeter and
a stopwatch over each 120 meter course.

The stationary firing was accomplished off vehicle in
the standing position. There were four shooters, each
equipped with an M16 Rifle, reflex sight and a stabilizer
and the necessary ancillary equipment.

Figures 2, 3 & 4 show the data accumulated during the
tests. These are averages of each 10 round firing group of
approximately four replications. After each firing, the
target was coded as to shooter, toward or across target,
reflex or weapon sight, stabilized or unstabilized. Figure
5 is a data sheet consisting of data taken from each target.
The data is in "cm" and gives the coordinates of the hit with
the origin at the center of the target. This particular
data sheet shows six hits in a ten round firing by shooter
#4 using a reflex sight stabilized when moving across the
target. The location of the hits in this case represented
data which was not used during moving firing analysis. It
was planned to measure mean radius for each ten round shot
group. This was impractical since the location of the missed
rounds was unknown, and any assumptions about the rounds that
missed would most certainly bias the results. Hit probability
was considered the most practical measure of performance. In
the stationary firing, all the rounds hit the target and
presented no problems in determining mean radius.

Figure 6 shows the factorial analysis printout for the
moving firing across target data. The main factors are S -
Sights, U - Stabilization, R - Speed, and G - Shooters. You
will notice that the F ratio for sights by shooter (SG) inter-
action is significant as well as the main factor, speed. The
significance level is .05 for all results. When an inter-
action is significant no inference can be made concerning
the main factors. The significant interaction indicates that
the main factors are not independent. We must investigate
the main effects by holding either sights or shooters con-
stant. Sights were held constant for this analysis. Figure
6 also shows the breakdown of the across target analysis into
separate sight analyses. The reflex sight indicates no
significant effects. The weapon sight shows the effects of
the three main factors significant. A covariance analysis of
reflex and weapon sights for each stabilization mode was
performed (Figure 7) to determine if a reduction in residual
mean square was significant. If so, covariance analysis would
indicate a linear regression (1 df, has to be linear if signif-
icant) of % hits versus vehicle speed. The reflex sight, based

upon the factorial analysis, indicated no significant
differences between stabilization modes, therefore the co-
variance analysis compared both stabilization modes simul-
taneously. There was a significant reduction in residual
due to regression and each shooter's slopes were statistically
parallel with the intercepts equal! This covariance analysis
concludes that a single regression line describes the per-
formance of the reflex sight stabilized or unstabilized.
The equation of the descriptive line is obtained using the
combined slope and the intercept. The examination of the
weapon sight using factorial analysis indicated a significant
difference between stabilization modes, speed and shooters,
therefore, the covariance analysis was performed examining
each stabilization level separately.

Figure 7 also shows the analysis for the weapon sight
stabilized and unstabilized across target firing. In the
stabilized mode there was a reduction in residual variance
due to regression; however, the slopes were not parallel and
the intercepts were not equal. This indicates that there was

a significant difference in performance of the shooters utilizing
the weapon sight stabilized. The results would seem difficult
to interpret, but it can be readily seen that in moving
across the target, the stabilizer would retard the angular
motion of the rifle as the shooter adjusts for the movement
of the vehicle and these results verify the difficulty the
shooters had in firing in this mode. The only reason this
is not apparent with the reflex sight stabilized is that the
number of points of alignment that must be made to hit the
target is 2 for the reflex sight and 4 for the weapon
sight and therefore it was probably easier to acquire the
target with the reflex sight despite the retardation of the
angular motion.

For the weapon sight unstabilized there is a reduction
in residual variance using covariance (i.e. there is linear
regression). The test for parallism shows no significant
difference indicating the slopes are parallel and the inter-
cept analysis shows intercepts equal.

This analysis concludes that a single regression line
describes the unstable firing results of the shooters when
traveling across the target. Figures 8, 9, and 10 illustrate
the results of the across target firing.

Reflex Sight: One linear regression describes reflex performance for stable and unstable and all shooters (Figure 8).

Weapon Sight:

Stable: Significant difference between shooters covariance analysis - linear regressions describe performance (Figure 9).

Unstable: No significant difference between shooters covariance analysis - one linear regression line describes weapon sight unstable performance for all shooters (Figure 10).

A comparison of reflex sight (stable and unstable) performance with weapon sight stabilized indicated no significant difference. A significant difference exists between reflex sight and weapon sight unstabilized. This is readily evident by comparing the levels of the regression lines. Figure 11 illustrates these comparisons - average line for each condition.

Figures 12 and 13 show the results of the moving vehicle toward target analysis.

1 - Factorial Analysis (Figure 12): There are several first and second order interactions which are significant, indicating that the main factors are not independent and we must go deeper to determine the causes of this significant interaction. The sights were examined separately.

2 - Reflex Sight: Again, a first order interaction (sights by shooters) is significant.

3 - Weapon sight analysis: Each of the main effects are significant.

4 - A covariance analysis of reflex sight (Figure 13) shows there is a slight reduction due to regression with the slopes parallel and intercepts equal. One regression line describes the performance independent of stabilization.

5 - The covariance analysis of weapon sight was examined for stabilized and unstabilized conditions separately and the results are also shown by Figure 13. In the stable toward target condition there is a slight reduction using covariance and the analysis shows slopes parallel and intercepts equal, so therefore a single regression line describes the results. However, in the unstable toward condition, the slopes are parallel but the intercepts are not equal. Therefore there is a significant effect caused by the shooters. Figures 14, 15, and 16 illustrate these results:

Figure 14 Reflex Sight: One linear presentation describes the results (stable or unstable).

Figure 15 Weapon Sight stabilized: One linear presentation describes the results.

Figure 16 Weapon Sight unstabilized: Shows the significant difference between shooters.

Figure 17 illustrates the comparison between reflex and weapon sights indicated that there was no difference between the reflex and weapon sights stabilized, and both the reflex sight and weapon sight stabilized are significantly better than the weapon sight unstabilized in the toward target mode.

The factorial analysis and covariance analysis for the stationary firing at three ranges was performed in the same manner as the toward and across target analysis. Figures 18, 19, and 20 summarize the results:

- There was a significant difference between pairs of gunners for the weapon sight stabilized (Figure 18) and unstabilized (Figure 19), and no significant difference in performance was noted due to stability. (Compare levels of Figure 18 and Figure 19).

- During stationary firing the reflex sight again exhibited no significant difference in performance due to

stability and there was no difference in the performance
of the shooters. Again, one single regression line describes
the performance.

- Figure 20 shows that the performance of the reflex
sight, stable and unstable, is significantly better (mean
radius lower) than the weapon sight, stable and unstable,
averaged performance.

CONCLUSIONS

This completed the analysis of the test data and
several interesting conclusions were drawn:

1 - The weapon sight performed significantly better
stabilized than unstabilized on moving vehicles.

2 - There was no difference in performance using a
stabilizer with weapon sights compared to weapon sights
unstabilized during stationary firing.

3 - Stability did not effect the performance of the reflex
sight.

4 - The reflex sight (stable or unstable) performed
significantly better than the weapon sight unstabilized,
and there was no difference detected between the reflex
sight (stable or unstable) and the weapon sight stabilized
during vehicle firing.

5 - There was a significant improvement in mean radius
using the reflex sight (stabilized or unstabilized) versus
the weapon sight stabilized or unstabilized.

6 - The reflex sight has a tendency to normalize the
shooters regardless of their ability. This normalization
appears to bring the level of ability toward the better
shooter. This normalization aspect is, of course, incon-
clusive since the levels of the shooters were unknown; how-
ever, further testing of the reflex sight to verify this
aspect was recommended. The effects of a stationary shooter
and a moving target with the stabilizer was not investigated;
however, based on the results of across target weapon sight

-73-

stabilized, it could be postulated that stabilization would affect the hit probability for this type of fire.

Based upon the results of this test and in conjunction with the USABAA, Frankford Arsenal is developing a reflex sight to perhaps replace the iron sights of the primary infantry weapon.

BIBLIOGRAPHY

Bowker, A. H. and G. J. Lieberman. Engineering Statistics. Englewood Cliffs, N.J: Prentice-Hall, 1959.

Cochran, W. G. and G. M. Cox. Experimental Designs. New York: John Wiley and Sons, 1959.
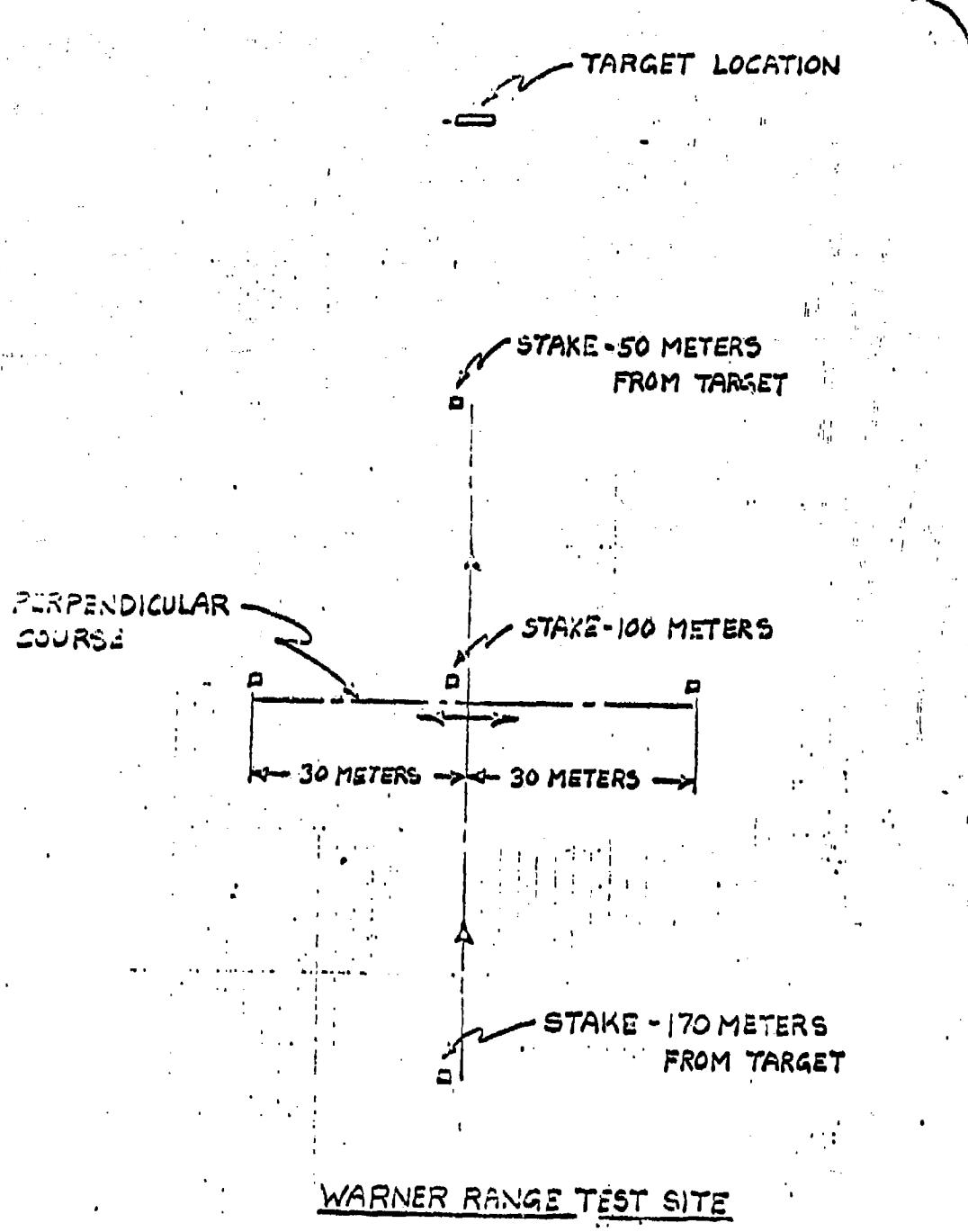
Jerome, C. R. Li. Statistical Inference. Ann Arbor, Michigan: Edward Brothers, Inc., 1964.

Kendall, M. G. and A. Stuart. Advanced Theory of Statistics Vol I and II. New York: Hafner Publishing Co, 1963.

Mood, A. M. and F. A. Graybill. Introduction to the Theory of Statistics, Second Edition. New York: McGraw-Hill, 1963.

Scheffé, H. The Analysis of Variance. New York: John Wiley and Sons, 1959.

Snedecor, G. W. Statistical Methods, Sixth Edition. Ames, Iowa: Iowa State University Press, 1967.

TARGET LOCATION

STAKE-50 METERS
FROM TARGET

PERPENDICULAR
COURSE

STAKE-100 METERS

|← 30 METERS →|← 30 METERS →|

STAKE - 170 METERS
FROM TARGET

WARNER RANGE TEST SITE

FIGURE 1

TOWARD TARGET (NO. OF HITS/10ROUNDS)

		REFLEX SIGHT		IRON SIGHT	
		STAB.	UNSTAB.	STAB.	UNSTAB.
	G_1	7.3	5.3	4.2	4.0
	G_2	6.5	5.7	7.0	5.5
8 M.P.H.	G_3	7.3	7.0	8.7	6.0
	G_4	5.2	5.7	7.0	5.5
	G_1	3.7	3.5	2.7	3.3
	G_2	4.0	4.5	4.7	4.3
10 M.P.H.	G_3	4.3	6.0	4.0	5.0
	G_4	3.0	4.0	6.0	4.0
	G_1	5.3	2.5	3.7	3.5
	G_2	4.7	3.0	4.7	2.5
12 M.P.H.	G_3	5.7	8.0	6.7	5.5
	G_4	6.0	7.0	4.2	4.0

FIGURE 2

ACROSS TARGET (NO. OF HITS/10 ROUNDS)

		REFLEX SIGHT		IRON SIGHT	
		STAB.	UNSTAB.	STAB.	UNSTAB.
	G_1	5.7	4.7	3.7	3.0
	G_2	6.0	4.0	6.8	6.0
8 M.P.H.	G_3	4.3	4.0	5.7	5.0
	G_4	4.5	4.0	4.2	5.5
	G_1	3.0	6.0	2.0	1.5
	G_2	5.0	4.5	4.7	3.5
10 M.P.H.	G_3	4.0	5.5	4.3	4.5
	G_4	4.5	3.5	5.0	1.5
	G_1	4.0	3.0	2.0	1.5
	G_2	3.3	4.5	3.7	1.5
12 M.P.H.	G_3	3.0	4.5	4.0	2.5
	G_4	3.7	2.5	5.0	3.0

FIGURE 3

STANDING - (MEAN RADIUS IN INCHES)

		REFLEX SIGHT		IRON SIGHT	
		STAB.	UNSTAB.	STAB.	UNSTAB.
50 M	G_1	6.65	3.41	3.76	4.46
	G_2	4.03	4.84	3.19	3.62
	G_3	5.43	3.45	3.31	3.68
	G_4	3.46	4.89	3.82	4.32
100 M	G_1	5.85	5.59	7.15	6.90
	G_2	5.81	4.89	5.12	5.75
	G_3	4.04	6.01	4.96	6.91
	G_4	5.34	7.14	7.52	7.91
150 M	G_1	7.35	7.17	9.22	11.11
	G_2	10.54	8.63	8.40	8.50
	G_3	6.07	8.18	7.26	8.31
	G_4	7.31	9.12	9.21	10.31

FIGURE 4

NO. _____46-4_____ STAB. _____Yes_____

SIGHT _____RFLX_____ LOCATION _____Handle:_____

RANGE _____100·_____ STATIONARY _____ MPH _12_

ROUND	X (HORIZ.)	Y. (VERT.)
1	.4.5	-39.5
2	19.8	-13.5
3	0.6	9.9
4	-47.4	9.5
5	-9.7	40.5
6	-12.7	80.7
7		
8		
9		
10		

ACROSS TARGET   FACTORIAL ANALYSIS

SOV	DF	SS	MS	F
1	1.	2.60333	2.60333	4.20719
2	1.	3.20333	3.20333	4.80750
3	2.	20.24291	10.12145	15.19000
4	3.	8.72916	2.90972	4.36685
5	1.	2.90083	2.90083	4.35351
6	2.	4.47041	2.23520	3.35455
7	3.	10.96666	3.65555	5.48619
8	2.	.45791	.22895	.34361
9	3.	3.36666	1.12222	1.68421
10	6.	3.64200	.60701	.91099
11	2.	4.43041	2.21520	3.32454
12	3.	.43593	.14527	.21803
13	6.	7.50700	1.25118	1.87774
14	6.	3.42450	.57076	.85659
15	6.	3.99791	.66631	.00000
TOTAL 16	47.	80.57916	.00000	.00000

FACTORIAL ANALYSIS

REFLEX SIGHT   ACROSS TARGET

SOV	DF	SS	MS	F
U	1	.0037	.0037	.024
R	2	5.5574	2.7787	2.711
G	3	1.9045	.6615	.645
UR	2	2.9576	1.4788	1.443
UG	3	2.8747	.9582	.935
RG	6	1.1693	.19408	.190
URG	6	6.1490	1.0248	
TOTAL	23	20.6962		

FACTORIAL ANALYSIS

WEAPON SIGHT

SOV	DF	SS	MS	F
U	1	6.1004	6.1004	6.83 *
R	2	19.1550	9.5779	10.73 *
G	3	17.7113	5.9038	6.61 *
UR	2	1.9309	.9654	1.
UG	3	.9279	.3093	
RG	6	5.8975	.9829	
URG	6	5.5553	.8926	1.10
TOTAL	23	57.6756		

FIGURE 6
80

SLOPES PARALLEL
INTERCEPTS EQUAL

ANALYSIS OF COVARIANCE

SLOPES NOT PARALLEL
INTERCEPTS NOT EQUAL

ANALYSIS OF COVARIANCE
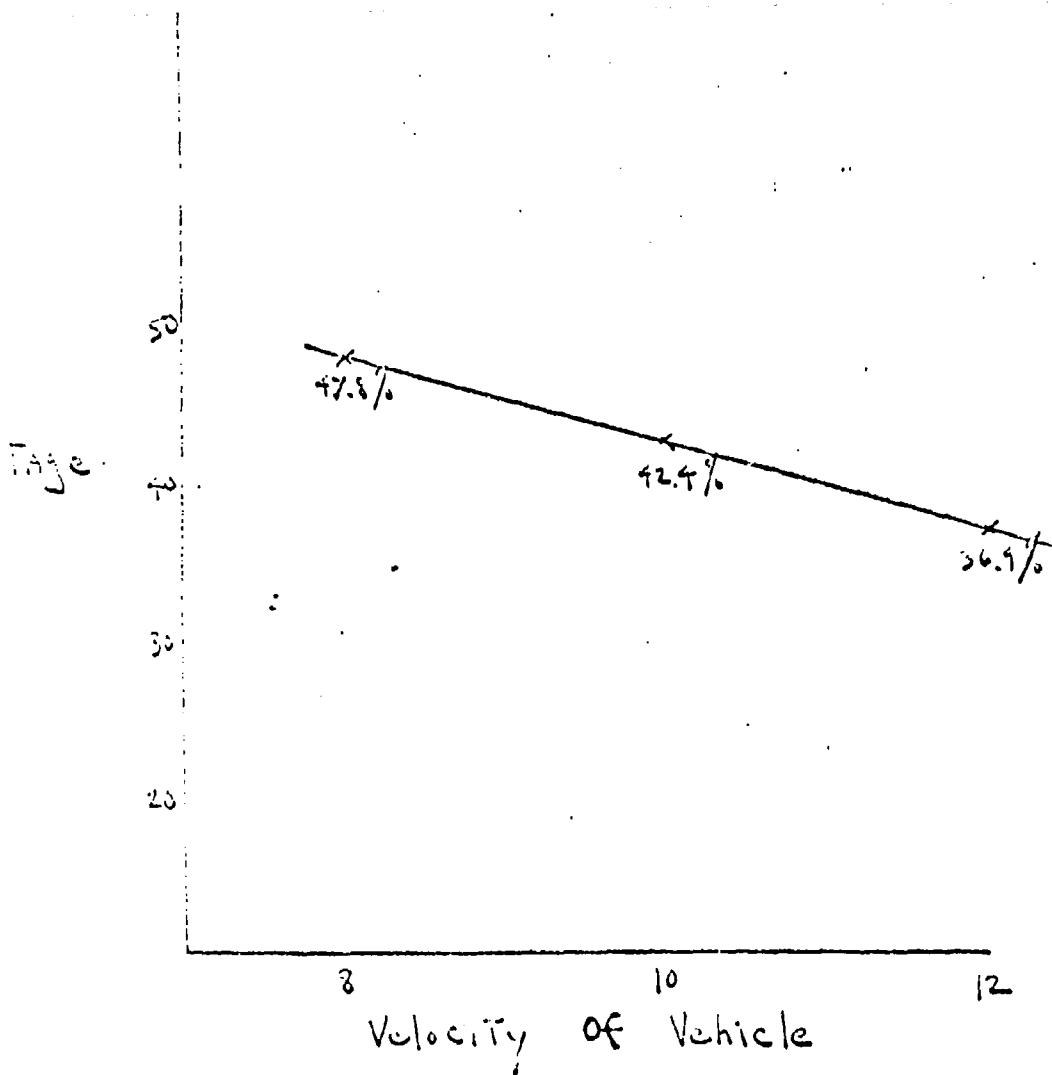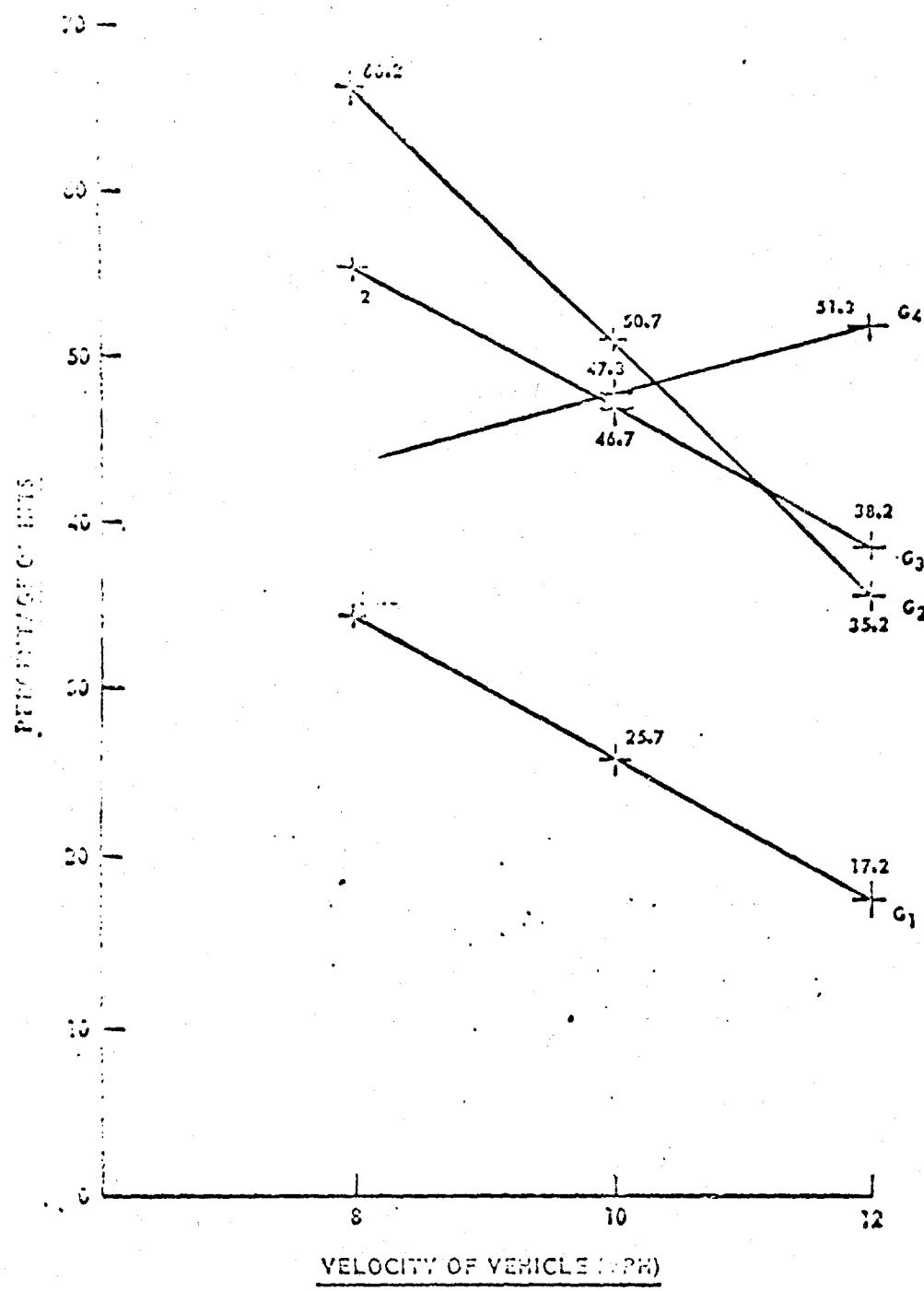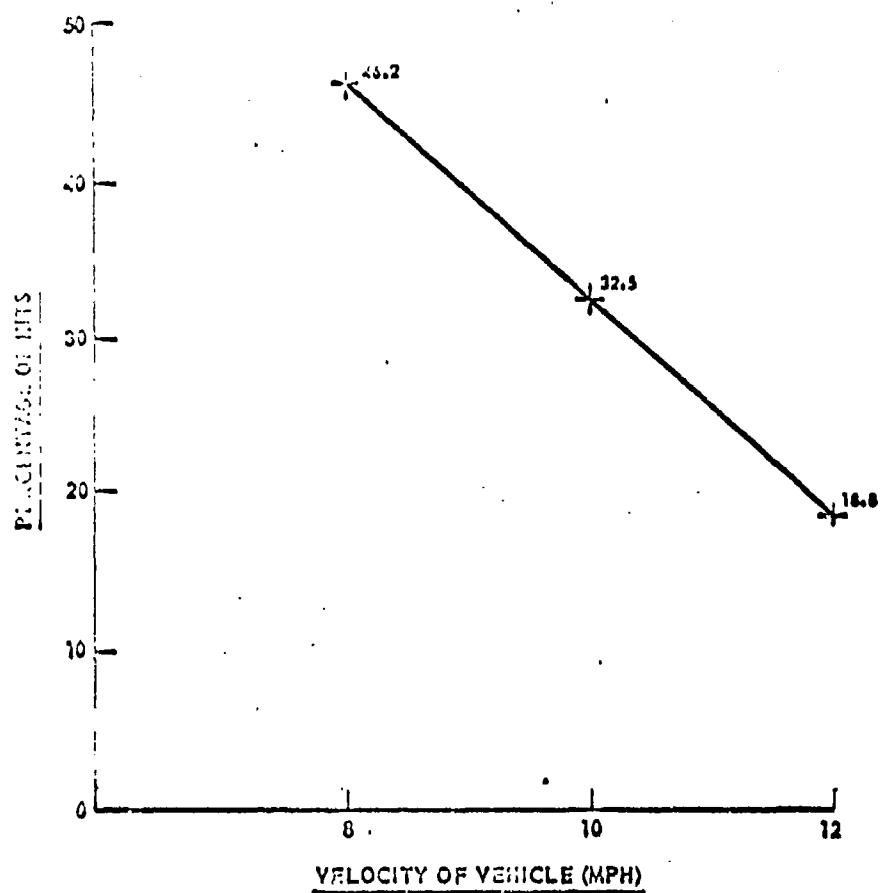
SLOPES PARALLEL
INTERCEPTS EQUAL

REFLEX SIGHT — ACROSS TARGET
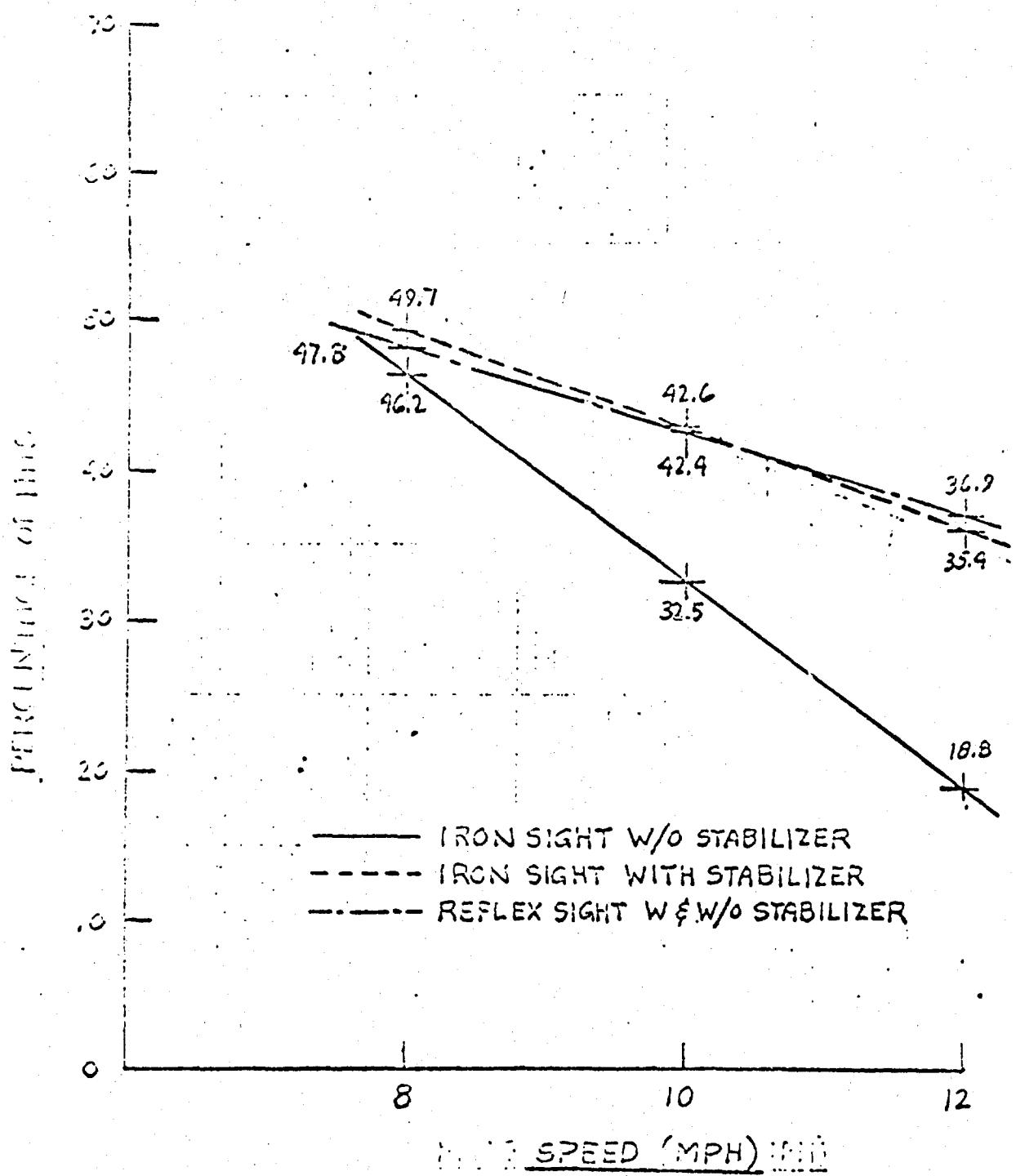
Weapon Sight With Stabilizer - M113 APC
Moving Across Target

FIGURE 9

Best Available Copy

Weapon Sight Without Stabilizer – M113 APC Moving Across Target

FIGURE 10

STANDARD "A" RIFLE TARGET (48" X 72")
M113 APC MOVING, PERPENDICULAR TO TARGET

FIGURE 11

SOV	D.	SS	MS	F
	1.	1.50520	1.50520	5.06541
	1.	2.66020	2.66020	8.95232 *
	2.	31.07541	15.53770	52.28861 *
	3.	27.72362	9.24107	31.10142 *
	1.	1.96020	1.96020	6.59663 *
	2.	2.25541	1.12770	3.79504
	3.	2.18895	.72965	2.45548 *
	2.	3.72041	1.86020	6.26010 *
	3.	2.53395	.84465	2.84248
	6.	7.25629	1.20937	4.06987
	3.	.06841	.02270	.07641 *
	3.	7.49395	2.49798	8.40640 *
	6.	4.74791	.79131	2.66300 *
	6.	9.16291	1.52715	5.13928 *
	6.	1.78291	.29715	.00000
	47.	106.11479	.00000	.00000

FACTORIAL ANALYSIS

TOWARD TARGET   RIFLE SIGHT

SOV	DF	SS	MS	F
U	1	.0266	.0266	--
A	2	18.1033	9.0516	19.33 *
C	3	11.8766	3.9589	8.45 *
UA	2	2.1234	1.0617	2.27
UC	3	8.1301	2.7100	5.78 *
AC	6	10.3434	1.7239	3.68
UAC	6	2.8099	.4683	
TOTAL	23	53.4133		

FACTORIAL ANALYSIS

TOWARD TARGET   WEAPON SIGHT

SOV	DF	SS	MS	F
U	1	4.5937	4.5937	7.40 *
A	2	15.2274	7.6137	12.30 *
C	3	18.0379	6.0126	9.70 *
UA	2	1.6426	.8213	1.32
UC	3	1.8979	.6326	1.02
AC	6	6.0759	1.0126	1.60
UAC	6	3.7203	.6201	
TOTAL	23	52.1962		

FIGURE 12

SOV	D.	SS	GS	F
		1.50520	.50520	5.06543
		2.66020	.66020	8.95232 *
		31.07641	15.53770	52.28861 *
		27.72102	9.24187	31.10142 *
		1.96020	1.96020	6.59633 *
		2.25541	1.12770	3.79564
		2.18895	.72965	2.45548
		3.72041	1.86020	6.26010 *
		2.53395	.84465	2.84248
		7.25625	1.20937	4.06987
		.04541	.02270	.07641
		7.49395	2.49798	8.40640 *
		4.74791	.79131	2.66300 *
		9.16291	1.52715	5.13928 *
		1.78291	.29715	.00000
	47.	106.11479	.00000	.00000

FACTORIAL ANALYSIS

TOWARD TARGET   REFLEX SIGHT

SOV	DF	SS	MS	F
	1	.0266	.0266	--
	2	18.1033	9.0516	19.33 *
	3	11.8766	3.9589	8.45 *
	2	2.1034	1.0517	2.27
	3	8.1301	2.7100	5.73 *
	6	10.3434	1.7239	3.68
	6	2.8099	.4683	
TOTAL	23	53.4133		

FACTORIAL ANALYSIS

TOWARD TARGET   WEAPON SIGHT

SOV	DF	SS	MS	F
	1	4.5937	4.5937	7.40 *
	2	15.2274	7.6137	12.30 *
	3	18.0378	6.0126	9.70 *
	2	1.6426	.8213	1.32
	3	1.8979	.6326	1.02
	6	6.0756	1.0126	1.60
	6	3.7206	.6201	
TOTAL	23	52.1962		

FIGURE 12

WEAPON SIGHT - STABLE TOWARD   ANALYSIS OF COVARIANCE

N-1	(Z)	(X2)	REDUCTION	(XY)	RESIDUAL	N-2	F

SUM
WITHIN
//-ISH

AMONG
WITHIN
TOTAL

SIGHT DEVIATION USING COVAR
SLOPES PARALLEL
INTERCEPTS EQUAL

WEAPON SIGHT UNSTABLE TOWARD TGT   ANALYSIS OF COVARIANCE

N-1	(Z)	(X2)	REDUCTION	(XY)	RESIDUAL	N-2	F

SUM
WITHIN
//-ISH

AMONG
WITHIN
TOTAL

SLOPES PARALLEL
INTERCEPTS NOT EQUAL

FIGURE 13

FIGURE 13

REFLEX SIGHT- TOWARD TARGET

Percentage
Hits

57.0%
52.2%
47.3%

8 10 12

Velocity of Vehicle
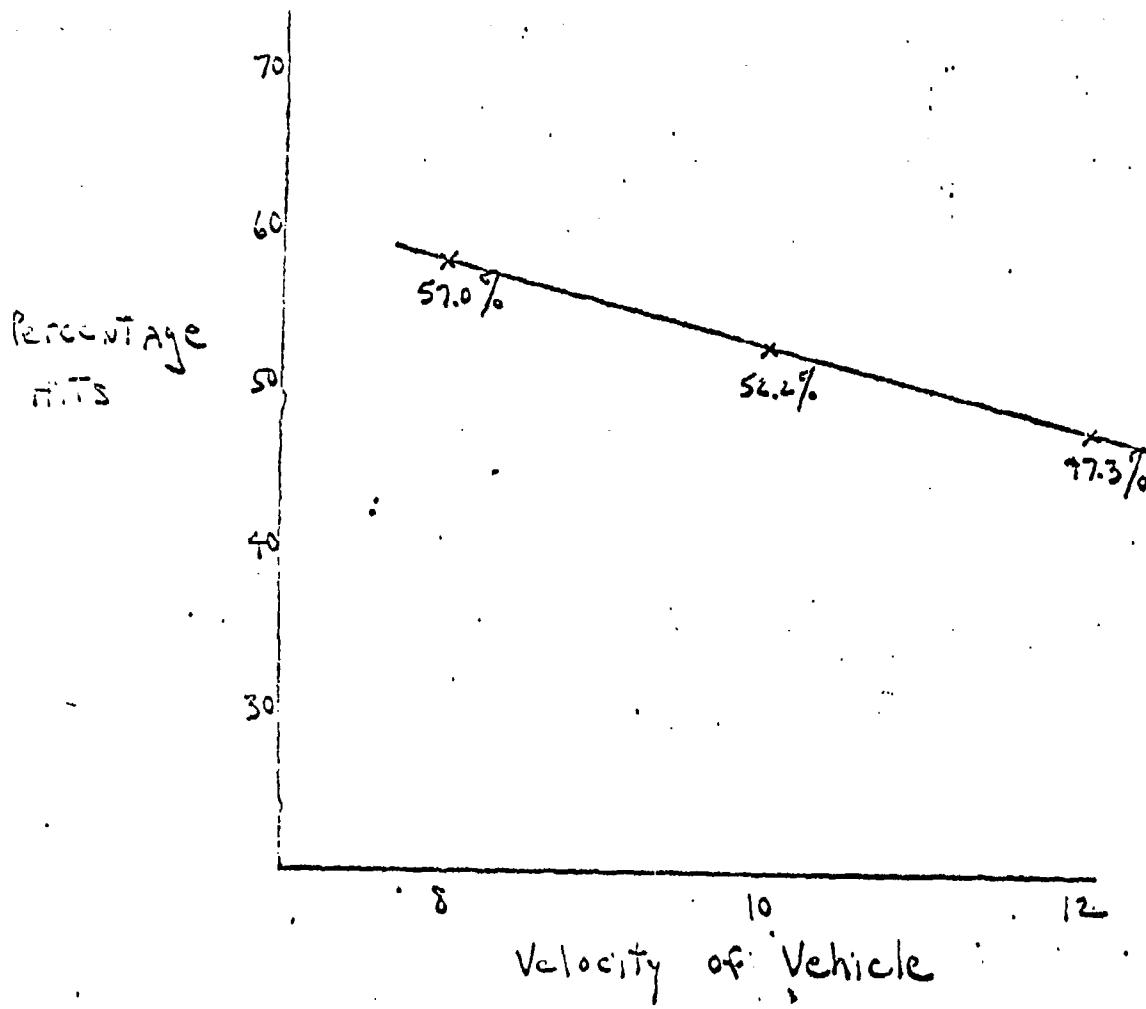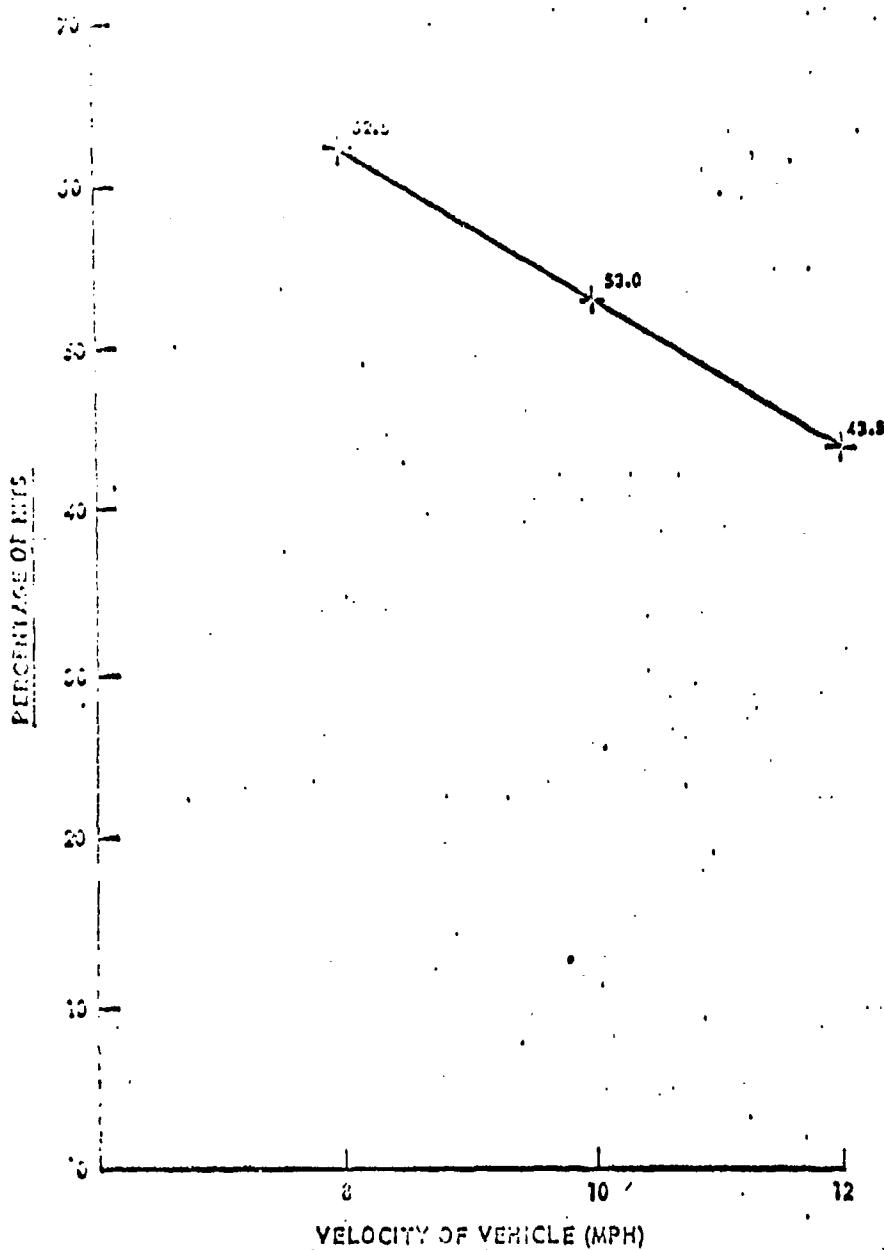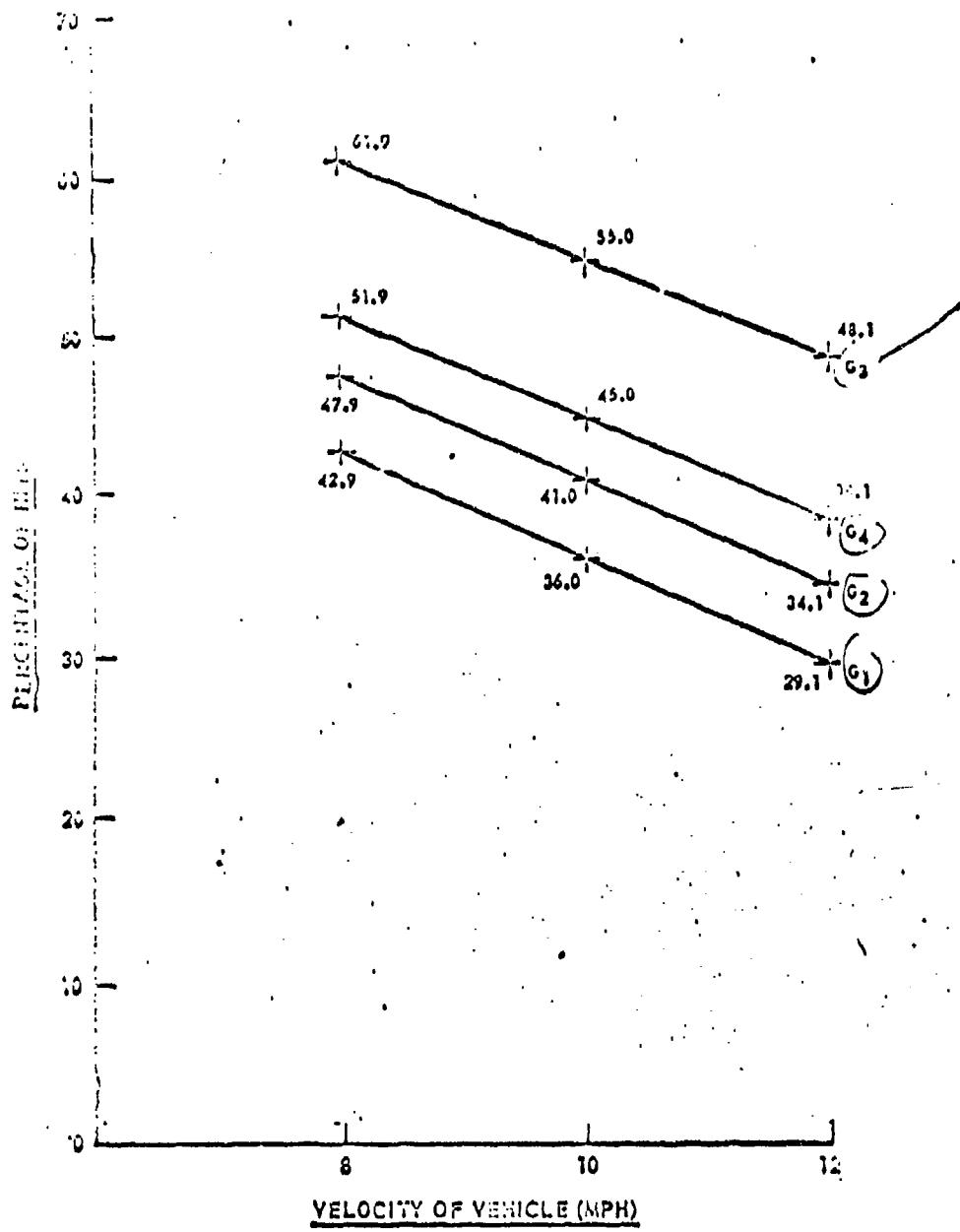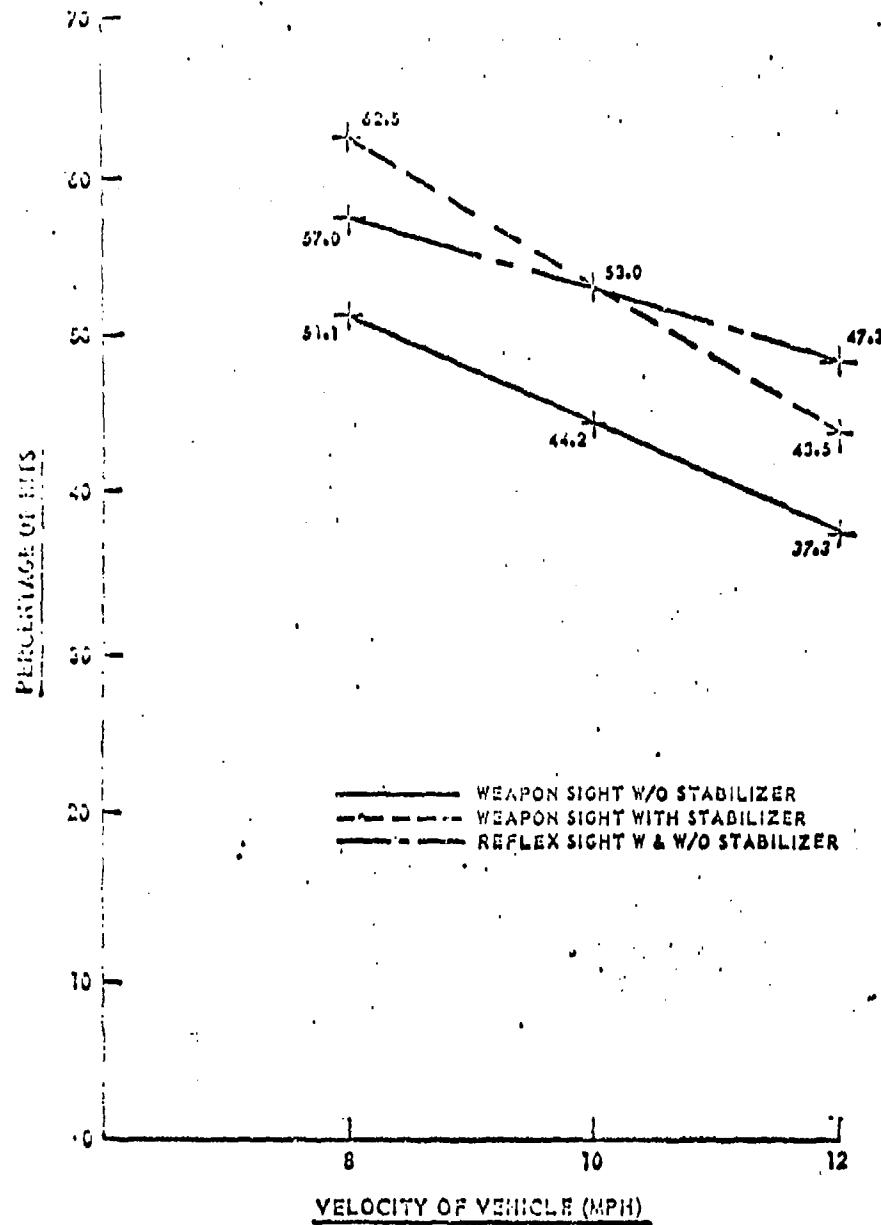
FIGURE 14
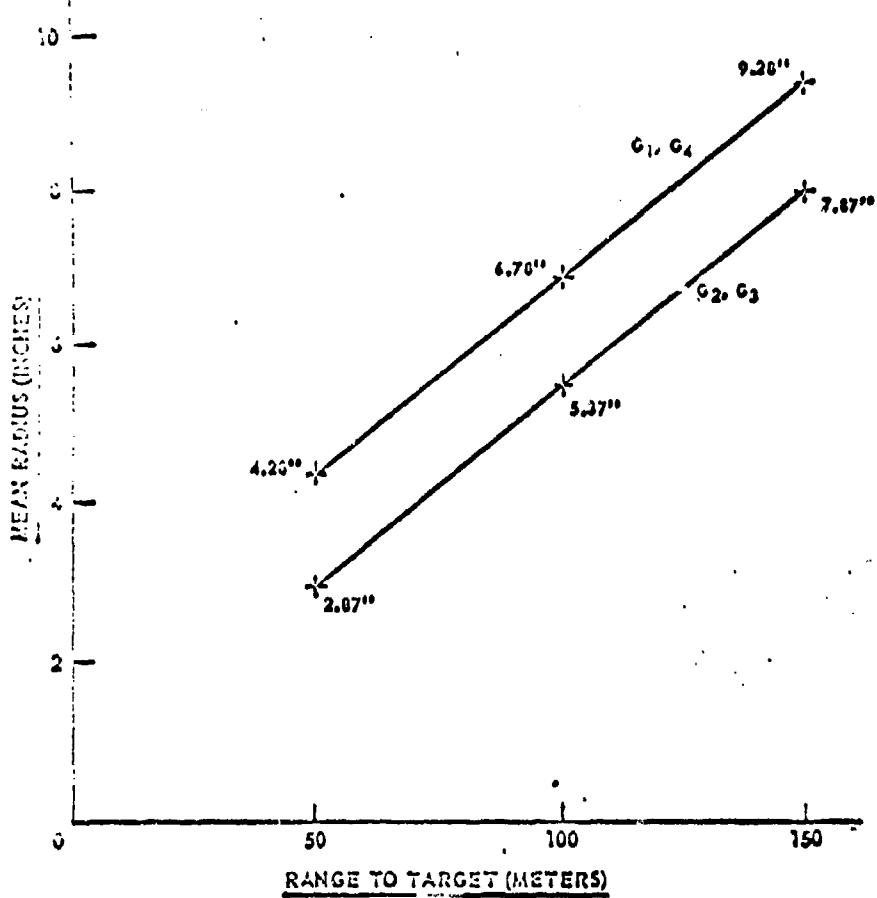
Weapon Sight with Stabilizer – M113 APC
Moving Towards Target
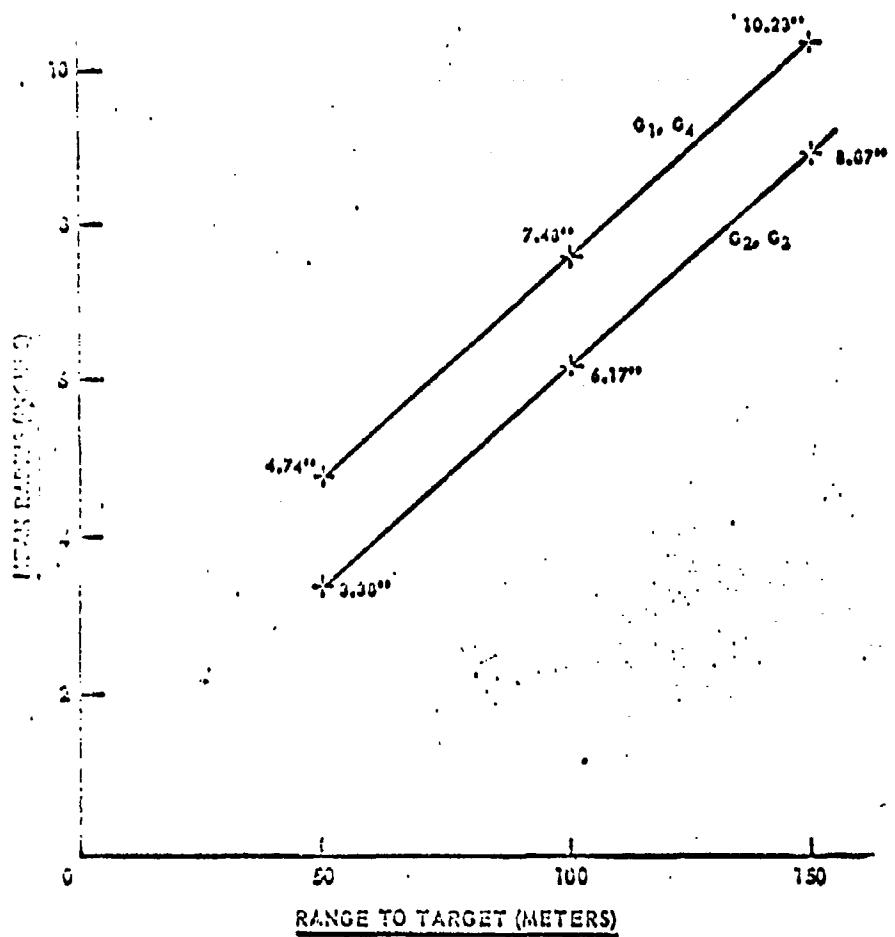
Standard "A" Rifle Target (48" x 72") – M113 APC Moving
Towards Target – Weapon Sight Without Stabilizer

Standard "A" Rifle Target (48" x 72") – M113 APC
Moving Towards Target
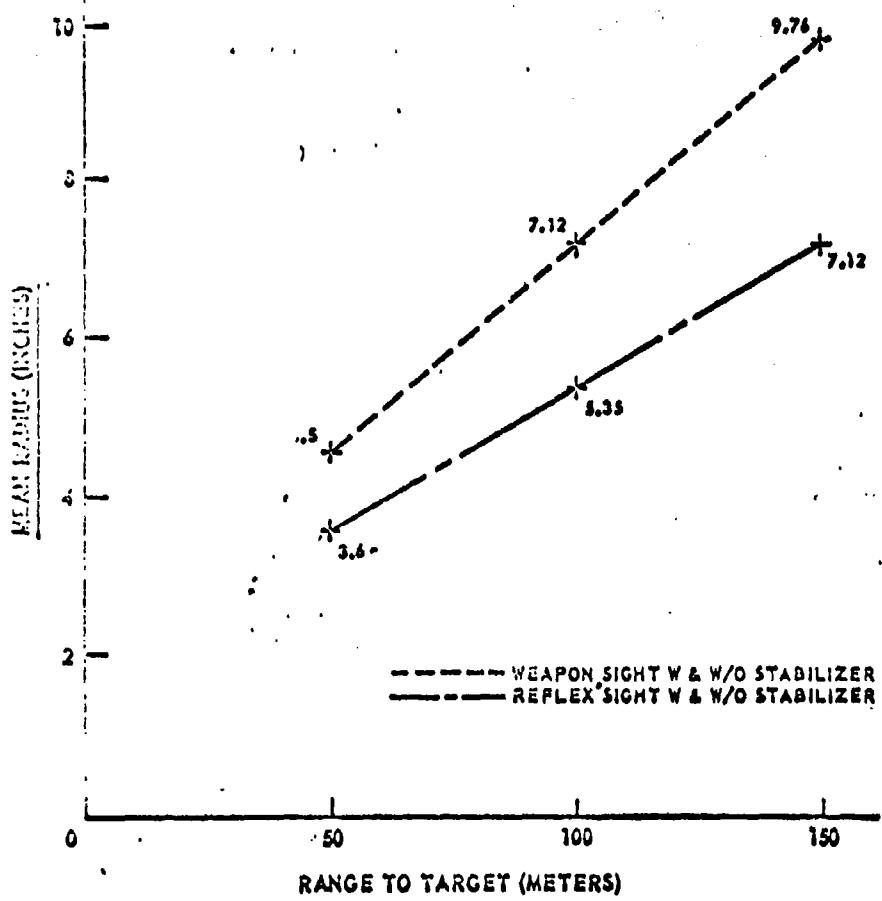
FIGURE 17

Weapon Sight - Stationary Firing with Stabilizer

FIGURE 18

-92-

Weapon Sight – Stationary Firing Without Stabilizer

FIGURE 19

Standard "A" Rifle Target (48" x 72") - Riflemen Standing

FIGURE 20

CHARACTERIZATION OF BALLISTIC EFFECTIVENESS
BY MAXIMAL TRAJECTORY

J. T. Wong and T. H. M. Hung
Systems Research Division
Research, Development and Engineering Directorate
HQ, US Army Weapons Command
Rock Island, Illinois

ABSTRACT

The purpose of this study is to introduce a measure by which the information content, possessed by the trajectory generated by a projectile, pertinent to the ballistic effectiveness of a weapon system can be characterized.

Having this concept and some appropriate mathematical tools at our disposal, a few relevant results are obtained. In particular, a characterization of the ballistic effectiveness of a weapon system in terms of the trajectory information is given.

The results obtained give some indications as to what conditions the data ensemble contains all the information offered by trajectories, concerning the weapon system under consideration. Also, a formula for the total of maximum recoverable amount of information possessed by a set of n trajectories (n shots) is given explicitly.

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

I. INTRODUCTION

A conventional method to obtain a set of terminal ballistic dispersion data experimentally for the purpose of analyzing the ballistic effectiveness of a weapon system is to fire the weapon upon a sufficiently large target plane. As the result, impact points are secured and upon which a well established analytical procedure is then applied to gain some concrete realization of the effectiveness of the weapon system under evaluation.

The purpose of this study is to establish a reasonable measure of information content in a given trajectory. Hence, such a measure can be applied to the data ensemble. Subsequently, on the basis of this concept and some other necessary mathematical notions, a few simple results are obtained.

The results so obtained give some indications as to what conditions the data ensemble contains all the information offered by a trajectory; in particular, are the data collected from the "plane" giving all the relevant information pertinent to the capability of the weapon system.

II. PRELIMINARIES

In this section we recall some of the well established concepts from the mathematical literatures. Consequently, the proofs of the cited theorems are omitted. Interested readers are referred to any standard mathematical analysis books such as Rudin [4], Dieudonne [1] and Dugundji [2].

Definition: Let R be the set of reals, and A, B be any two non-empty subsets of R^3 (the Euclidean 3-space). The distance between A

and B is

$$d(A,B) = \inf \{\rho(a,b) \mid a\varepsilon A, b\varepsilon B\}$$

$$= \inf_{\substack{a\varepsilon A \\ b\varepsilon B}} \rho(a,b)$$

where "inf" of C is the greatest lower bound of a non-empty set C and ρ denotes the usual Euclidean distance function on R^3; i.e.,

$$\rho: \ R^3 \times R^3 \to [0,\infty)$$

defined by

$$\rho(x,y) = \sqrt{\sum_{i=1}^{3} (x_i - y_i)^2} \ , \quad x,y\varepsilon R^3$$

where, for example, x_i is the ith component of x.

If one of the sets A,B is the null set ϕ, then

$$d(A,B) = \infty \ .$$

Theorem 2.1: Let A,B be non-empty closed subsets in R^3. Then there exists $a_0\varepsilon A$ and $b_0\varepsilon B$ such that

$$d(A,B) = \rho(a_0, b_0)$$

Theorem 2.2: Let f and g be two continuous mappings on a metric space E into a metric space E'. The set A of all points $x\varepsilon E$ such that $f(x) = g(x)$ is closed in E.

Theorem 2.3: If A,B and C are subsets of R^3 such that $B\varepsilon C$, then

$$d(A,B) \geq d(A,C).$$

<u>Definition</u>: For every $i \epsilon \Omega_n$, let ψ_i be a continuous mapping on R^3 into R with respect to the usual distance functions. Then a non-empty subset of points $s = (s_1, s_2, s_3)$ of R^3 is called a surface S if

1) $S = \bigcup_{i=1}^{n} S_i$

ii) $S_i = \{ s \epsilon R^3 | \psi_i(s) = k_i , k_i \epsilon R \}.$

The definition for a surface defined here is different from that which one may find in mathematical literature, for which $n = 1$ and ψ_i is differentiable, [4, p. 209], for example.

Now we have the necessary mathematical tools to introduce some new notions and proceed as follows.

In the course of evaluating experimentally the terminal ballistic effectiveness or capability of a weapon system (particularly small arms) in delivering a projectile to a preassigned fixed "point" (aiming point), we normally collect the terminal dispersion data by placing an aiming point on a sufficiently large plane upon which the weapon system is being fired, Fallin [3]. In this case, the dispersion data are the set of all impact points on the plane, with respect to a fixed two-dimensional cartesian coordinate system.

Every impact point on the plane possesses three intrinsic properties of the corresponding trajectory. Obviously, the three properties are magnitude, direction, and the point, which is itself a point on the trajectory. The vector properties of an impact point provide not only a

-99-

measure of closeness of a given trajectory to the aiming point but also the bias of the weapon system.

As for the plane, its function is to pick a point on the trajectory, for without such a plane we do not know in general the vector character of a point on a trajectory. In view of this observation, we could use a surface to assume the function of a plane as well. Since every plane is a surface according to the definition above, it would be interesting to find out what kind of surfaces would be best to use instead of a plane. The word "best" calls for some criteria with which a comparison can be made among all the admissible surfaces. The necessity of an appropriate quality measure **provides** part of the motivations for defining the following concepts.

The foregoing discussions imply that, for each point on a trajectory, there corresponds some measurable quantities indicating the capability of a weapon system in delivering a projectile to the aiming point. A point, for example, on a trajectory possesses some information concerning closeness to the aiming point -- consequently, the accuracy of the system. Obviously, the correspondence between each point on a trajectory and the corresponding information possessed by that point gives in a natural way a function between two sets. Such a function may be defined as follows:

Definition: Let F be a trajectory. A real-valued function I on F into the reals is defined by

$$I(f) = \frac{1}{1 + \rho(f,\theta)} \ , \quad f \epsilon F$$

where ρ is the usual distance function on the euclidean three-space, and θ is the aiming point and, for convenience, is taken to be the origin of the coordinate system.

We observe that the function I is well-defined since ρ is. Also, the continuity of I is implied by that of $\rho(\ ,\theta)$ on a subset of R^3. Furthermore, since ρ is a non-negative real-valued function, it follows that $0 \leq I(f) \leq 1$ for any point f belongs to a trajectory F.

Physically, the quantity $I(f)$ is a measure of closeness of a point f on the trajectory F as related to the aiming point. Consequently, it contains knowledge concerning the accuracy of the weapon system under consideration. For these reasons, we may interpret the quantity $I(f)$ as the amount of information associated with F at the point f. The maximum attainable amount of information possessed by a trajectory F is given by

$$\underset{f \epsilon F}{\text{Sup}}\ I(f) = \frac{1}{1 + \underset{f \epsilon F}{\text{inf}}\ \rho(f, \theta)}$$

Since F is closed and bounded -- the direct image of a continuous mapping on a compact set is compact, we have

$$\underset{f \epsilon F}{\text{Sup}}\ I(f) = \underset{f \epsilon F}{\text{max}}\ I(f) = \frac{1}{1 + \underset{f \epsilon F}{\text{min}}\ (f, \theta)}$$

Therefore, the maximum amount of information possessed by F is recoverable at some point $f_0 \epsilon F$. Symbolically, $\underset{f \epsilon F}{\text{Max}}\ I(f) = I(f_0)$ (The existence of such a point f_0 is asserted by Theorem 2.1.) As for the uniqueness, we make the following stipulation:

Assumption: There exists at most one point $f_0\epsilon F$ such that $I(f_0)$ is the maximum recoverable amount of information of F.

The assumption so stated is not indispensable. However, as a result of this stipulation, subsequent analysis is simplified. In fact, the assumption is reasonable, for in the neighborhood of the aiming point, the radius of curvature of a physically realizable trajectory F at f is greater than $\rho(f_0,\theta)$, if $f \neq f_0$; that is, F is tangent to the sphere

$$\{x\epsilon R^3 | \rho^2(x,\theta) = \rho^2(f_0,\theta)\}$$

at $f_0\epsilon F$.

In view of the foregoing remarks, for each trajectory F there exists one and only one point f_0 belonging to F such that I assumes its maximum value at f_0.

Before defining the last concept for this analysis, we recall for a trajectory F with $g\epsilon F$, $I(g)$ is the amount of information assigned to F at g. Equivalently, $I(g)$ is the amount of information gained about F by observing (sampling) g. Having this interpretation in mind, we make the following definition.

Definition: Let F be a trajectory. A real valued function $\overline{I}$ on F into the unit interval is defined by the formula

$$\overline{I}(g) = \max_{f\epsilon F} I(f) - I(g) \, , \, g\epsilon F$$

We note that $0 \leq \overline{I}(g) \leq I(f_0)$ and $\overline{I}(g) = 0$ if and only if $g = f_0$. It is clear that $\overline{I}(g)$ is the amount of information lost concerning F due to observing g.

Remark: If, in the definition of a trajectory, t_0 is taken to be the time at which the projectile leaves the muzzle point of a weapon, then the muzzle velocity $\dot{f}(t_0)$ is a well defined quantity. Hence, f is differentiable at t_0 also. Furthermore, the definition does not account for the case in which a round is misfired due to, for instance, malfunction of the weapon. In such an event, the trajectory is the null set in R^3.

IV. MAIN RESULTS

With the foregoing concepts at our disposal, we may deduce a few trivial consequences. The proof of the first theorem is immediate and is omitted.

Theorem 4.1: Let $\{F_i | i \in \Omega_n\}$ be a set of trajectories and, for each $i \in \Omega_n$, f_{i0} be the unique point of F_i for which $\max_{f \in F_i} I(f) = I(f_{i0})$. Then

$$\sum_{i=1}^{n} I(f_{i0}) = n$$

if and only if everyone of the n trajectories hits the aiming point $\hat{o}$.

Before we proceed to the second theorem, we need the following fact at our disposal.

Lemma: Every surface is a closed set.

Proof: By definition a surface S is a non-empty subset of R^3 such that

$$S = \bigcup_{i=1}^{n} S_i, \quad \text{for some positive integer } n, \text{ and for } i \in \Omega_n$$

$$S_i = \{s \in R^3 | \psi_i(s) = k_i; \quad k_i \in R\}.$$

Now by definition V_i is a continuous function on R^3 into R, and the constant function C_i defined by

$$C_i(s) = k_i \, , \quad s \varepsilon R^3$$

is also a continuous mapping on R^3 into R. By Theorem 2.2, the set

$$\{s \varepsilon R^3 | V_i(s) = C_i(s)\}$$

is closed and is precisely S_i, for $C_i(s) = k_i$. The assertion follows from the fact that a finite union of closed set is closed.

Theorem 4.2: Let $\{F_i | i \varepsilon \Omega_n\}$ be the set of trajectories corresponding to the n rounds fired from a weapon system at an aiming point, and S be a surface. Furthermore, let A_i be defined by

$$A_i = \{g \varepsilon R^3 | g \varepsilon S \cap F_i\}.$$

Then

$$\sum_{i=1}^{n} \max_{g \varepsilon A_i} I(g) = \sum_{i=1}^{n} \max_{g \varepsilon F_i} I(g)$$

is a necessary and sufficient condition for no information is being lost by the experiment.

Proof: Sufficiency

$$A_i \subset F_i \quad \text{implies} \quad \inf_{g \varepsilon F_i} \rho(g, \theta) \leq \inf_{g \varepsilon A_i} \rho(g, \theta) \quad \text{by Theorem 2.3.}$$

Since $\rho(g, \theta) \geq 0$. We have

$$\frac{1}{1 + \inf_{g \varepsilon A_i} \rho(g, \theta)} \leq \frac{1}{1 + \inf_{g \varepsilon F_i} \rho(g, \theta)} \, ,$$

-104-

or equivalently

$$\sup_{g \in A_1} I(g) \leq \sup_{g \in F_1} I(g).$$

Therefore, for each $i \in \Omega_n$

$$\sup_{g \in F_1} I(g) - \sup_{g \in A_1} I(g) \geq 0.$$

By the lemma above, S is closed and so is F_1. It follows that $A_1 = S \cap F_1$ is closed. Hence, supremum is the same as maximum. We have

$$\max_{g \in F_1} I(g) - \max_{g \in A_1} I(g) \geq 0.$$

Now, the hypothesis implies that for each $i \in \Omega_n$

$$\max_{g \in F_1} I(g) - \max_{g \in A_1} I(g) = 0.$$

If $A_1 \neq \phi$ for each $i \in \Omega_n$, then there exists $g_{10} \in A_1$ such that

$$I(g_{10}) = \max_{g \in A_1} I(g)$$

Therefore

$$\overline{I}(g_{10}) = \max_{g \in F_1} I(g) - I(g_{10}) = 0.$$

Thus, the amount of information lost concerning F_1 due to observing g_{10} is zero, and the sum

$$\sum_{i=1}^{n} \overline{I}(g_{i0}) = 0$$

vanishes asserts that no information is being lost by the experiment.

Now suppose for some $i \epsilon \Omega_n$, $A_i = \phi$. Then

$$\max_{g \epsilon A_i} I(g) = \sup_{g \epsilon A_i} I(g) = 0$$

The first equality follows from the fact that the null set ϕ is closed and the convention of the supremum over the null set yields the last equality. So for this i,

$$\max_{g \epsilon F_i} I(g) - \max_{g \epsilon A_i} I(g) = \max_{g \epsilon F_i} I(g) > 0$$

The strict inequality follows from a remark in section 5 that $F_i \neq \phi$. Thus if $A_i = \phi$ for some $i \epsilon \Omega_n$, then the hypothesis does not hold. We may then deduce that the conclusion is true.

It remains to prove the necessary condition. No information is lost and the fact that $F_i \neq \phi$, $i \epsilon \Omega_n$, imply $A_i \neq \phi$. Then there exists $f_{i0} \epsilon A_i$ such that

$$\max_{g \epsilon A_i} I(g) = I(f_{i0}).$$

Also

$$0 = \max_{g \epsilon F_i} I(g) - I(f_{i0}) = \max_{g \epsilon F_i} I(g) - \max_{g \epsilon A_i} I(g)$$

Summing over the last term on i gives the assertion, and the proof of the theorem is complete.

As the immediate consequences of this result, we have

Corollary 4.3: Suppose A_i for each $i \in \Omega_n$ contains one and only one point. There exists at least one $i \in \Omega_n$ such that, for $g_i \in A_i$, $I(g_i) > 0$ if and only if

$$\sum_{i=1}^{n} \max_{f \in F_i} I(f) > \sum_{i=1}^{n} I(g_i).$$

Corollary 4.4: Let S and S^* be two surfaces such that $F_i \cap S$ is a proper subset of $F_i \cap S^*$ for at least one i. Then, for a set of n trajectories,

$$\sum_{i=1}^{n} \max_{f \in F_i \cap S} I(f) \leq \sum_{i=1}^{n} \max_{f \in F_i \cap S^*} I(f) \leq \sum_{i=1}^{n} \max_{f \in F_i} I(f)$$

Theorem 4.5   Let F be a trajectory and S be a surface. Let A be the intersection of F and S. Suppose

$$\max_{f \in A} I(f) < \max_{f \in F} I(f)$$

Then there exists a sequence $(h_i)_{i=1}^{\infty} \subset F$ such that

i) $\displaystyle \max_{f \in F} I(f) = \lim_{i \to \infty} I(h_i)$

ii) there is a positive integer N such that, for all

$i \geq N$, $I(h_i) > \max_{f \in A} I(f)$

Proof: For $F \neq \emptyset$, Theorem 2.1 asserts that there is a point $f_0 \in F$ such that

$$\max_{f \in F} I(f) = I(f_0).$$

Since F is closed, then there exists a sequence $(h_i)_{i=1}^{\infty} \subset F$ such that $\lim_{i \to \infty} h_i = f_0$. Also we have seen in section 3 that the function I is continuous; it is this property for which the following first equality holds.

$$\lim_{i \to \infty} I(h_i) = I(\lim_{i \to \infty} h_i) = I(f_0) = \max_{f \in F} I(f).$$

It remains to be shown the last assertion of the theorem. If $A = \emptyset$, ii) holds trivially, since $F \neq \emptyset$ and we may pick $N = 1$. Suppose $A \neq \emptyset$. Again, Theorem 2.1 implies there is a point $g_0 \in A$ such that

$$\max_{f \in A} I(f) = I(g_0).$$

Now, by hypothesis we obtain $\rho(f_0, \theta) < \rho(g_0, \theta)$. Also we may pick an $\varepsilon > 0$ so that $\rho(f_0, \theta) + \varepsilon < \rho(g_0, \theta)$. For this ε and the fact that

$$\lim_{i \to \infty} h_i = f_0$$

We conclude that there exists a positive integer N such that for all $i \geq N$

$$\rho(h_i, f_0) < \varepsilon.$$

Using these facts and the well-known properties of the metric ρ, we obtain

$$\rho(h_i, \theta) \leq \rho(h_i, f_0) + \rho(f_0, \theta) < \varepsilon + \rho(f_0, \theta) < \rho(g_0, \theta).$$

-108-

Proof: For $F \neq \Phi$ Theorem 2.1 asserts that there is a point $f_0 \epsilon F$ such that

$$\max_{f \epsilon F} I(f) = I(f_0).$$

Since F is closed, then there exists a sequence $\{h_i\}_{i=1}^{\infty} \subset F$ such that $\lim_{i \to \infty} h_i = f_0$. Also we have seen in section 3 that the function I is continuous; it is this property for which the following first equality holds.

$$\lim_{i \to \infty} I(h_i) = I(\lim_{i \to \infty} h_i) = I(f_0) = \max_{f \epsilon F} I(f).$$

It remains to be shown the last assertion of the theorem. If $A = \phi$, ii) holds trivially, since $F \neq \phi$ and we may pick $N = 1$. Suppose $A \neq \phi$. Again, Theorem 2.1 implies there is a point $g_0 \epsilon A$ such that

$$\max_{f \epsilon A} I(f) = I(g_0).$$

Now, by hypothesis we obtain $\rho(f_0, \theta) < \rho(g_0, \theta)$. Also we may pick an $\epsilon > 0$ so that $\rho(f_0, \theta) + \epsilon < \rho(g_0, \theta)$. For this ϵ and the fact that

$$\lim_{i \to \infty} h_i = f_0$$

We conclude that there exists a positive integer N such that for all $i \geq N$

$$\rho(h_i, f_0) < \epsilon,$$

Using these facts and the well-known properties of the metric ρ, we obtain

$$\rho(h_i, \theta) < \rho(h_i, f_0) + \rho(f_0, \theta) < \epsilon + \rho(f_0, \theta) < \rho(g_0, \theta).$$

Obviously, $1 + \rho(h_i, \theta)$, $1 + \rho(g_0, \theta) > 0$. It follows that

$$\frac{1}{1 + \rho(h_i, \theta)} > \frac{1}{1 + \rho(g_0, \theta)} = I(g_0) = \max_{f \in A} I(f)$$

for all $i \geq N$, where N depends on ε.

Remark: We recall that in section 3 the definition of a surface requires every surface being non-empty. The restriction does not impose any further constraint on the results in this section; that is, the theorems and their consequences remain valid for the case in which S is empty and in particular the lemma remains true since the null set is closed. A trivial example of a null surface is given by

$$S = \{x \in R^3 | \rho^2(ax, bx_0) = -1, \ a, \ b \in R, \ x_0 \in R^3\}.$$

V. CONCLUSIONS

We conclude the study by stating explicitly some of the observations deduced from the foregoing results.

For a given experiment,

$$\sum_{i=1}^{n} \max_{f \in F_i} I(f)$$

is the total of the maximum recoverable amount of information possessed by a set of n trajectories. Theorem 4.1 states that such a total is equal to n if and only if the weapon system gives a perfect performance concerning its capability in delivering the n projectiles to the aiming point.

Theorem 4.2 asserts that, in the course of evaluating the terminal ballistic effectiveness of a weapon based on the results of analyzing the terminal dispersion data collected from the surface, an "ideal surface" to be used is one which perserves all the information offered by the experiment; equivalently, an ideal surface is one having a set of n trajectories as its subset. A simple example of such a surface S_I is

$$S_I = \bigcup_{\lambda \in [\alpha, \beta]} \{(x_1, x_2, x_3) \in R^3 \mid x_1 = \lambda; \ x_2, \ x_3 \in R\}$$

where

$$\alpha = \min \Lambda_0$$
$$\beta = \max \Lambda_1$$
$$\Lambda_j = \{f_i(t_j) \mid f_i \in F_i, \ i \in \Omega_n\} , \ j = 0,1$$

According to our definition of a surface, S_I is not a surface, for S_I is a union of uncountably many planes. Therefore, a perfect surface does not exist. Also, the second consequence of the theorems tells us that there is a positive information gained by extending the surface being used in such a way there is an increase in the data ensemble for at least one trajectory.

Finally, the last result concludes that there is a sequence of points in a given trajectory such that the limit of the image of the sequence under I *is precisely* the maximum information possessed by the trajectory. Also, if there is a loss of information, such a loss can be reduced by using an algorithm to recover some of the loss.

Remark: In view of S_I defined above, a reasonable surface to be used in an experiment is one which consists of $2m+1$ sufficiently large planes having the aiming point placed on the center of one plane and the remaining $2m$ planes erected at equidistance from each other on both sides of the center plane.

VI. REFERENCES

1. Dieudonne, J., _Foundations of Modern Analysis_, Academic Press, New York and London, 1960.

2. Dugundji, James, _Topology_, Allyn and Bacon, Inc., Boston, 1966.

3. Fallin, H. K., "Analysis of Machine-Gun Burst Dispersion Data With Corresponding Effectiveness Model", AMSAA TN-33, Aberdeen Proving Ground, Maryland, July 1969.

4. Rudin, W., _Principle of Mathematical Analysis_, McGraw-Hill Book Company, New York, 1964.

AN EFFECTIVENESS MODEL FOR BURST FIRES

ON VOLUME TARGETS

T. H. M. Hung and J. T. Wong
Systems Research Division
Research, Development and Engineering Directorate
HQ, US Army Weapons Command
Rock Island, Illinois

ABSTRACT

On the basis of three-trivariate normal distributions, an
effectiveness model for analyzing a machine gun firing bursts on volume
targets is developed. The development is essentially an extension of
an existing model based on three-bivariate normal distributions.
        A regular parallelpiped target is employed. The effectiveness
criteria, the probability of incapacitation and the expected number of
hits per burst, are expressed as functions of hit probabilities. Finally,
some numerical examples and observations are given.

The remainder of this paper was reproduced photographically from the
manuscript submitted by the author.

Preceding page blank

1. INTRODUCTION

Several studies have been conducted for analyzing effectiveness of a machine gun firing bursts upon a specified target [1,2,3,4]. In these studies effectiveness is defined in terms of two measurable parameters: $E(H)$ and $P(I)$, the expected number of hits and the probability of incapacitation per burst, respectively. On the basis of these effectiveness criteria, the foregoing investigations have rendered some basic models for effectiveness analysis by using a two, three or four-bivariate normal distributions. The models so obtained have provided a base for analyzing effectiveness on area targets.

In this report a three-trivariate noral distribution model is developed for three dimensional target-volume targets. The development is essentially an extension of the work in [3]. Thus, no further physical justification for the modeling is rendered in this endeavor.

As was known in [3] the dispersion of the impact points from a burst on area target revealed some interesting measurable phenomena. In particular, the dispersion data showed that the distribution of the first impact point of a given burst differentiated itself significantly from that rendered by the subsequent projectiles. For this reason, it was assumed that the behavior of these random phenomena can be characterized by two random vectors X^1 and X^2 respectively. Furthermore, the offset of the mean of the center of impact of the subsequent projectile from that of the first constituted the third random vector, denoted by X^3. As the essential

premises to be used for modeling, it is assumed that these observable facts can be carried over to the third dimension and each random vector has a mutually independent, trivariate normal distribution.

On the strength of the foregoing stipulations, the problem is solved for the case in which the configuration of the target is a regular parallelepiped. In the formulations of the effectiveness parameters, $E(H)$ and $P(I)$, are attained by expressing these parameters as functions of $P(F)$ and $P(S)$ - the hit probability of the first round and that of the subsequent ones. Finally, numerical examples as well as some interesting observations are given.

2. STATEMENT OF THE PROBLEM

2.1 Modeling

Referring to Fig. 1, let us denote a fixed point (x_1^j, x_2^j, x_3^j) in E^3, the euclidean 3-space, by x^j, and the differential of x^j in E^3 to be dx_1^j, dx_2^j, dx_3^j ; i.e., $dx^j = dx_1^j \cdot dx_2^j \cdot dx_3^j$. A solid parallelepiped target, as shown by the dotted lines, is located in E^3 with its surfaces intersecting the coordinates of $x^j = (x_1^j, x_2^j, x_3^j)$ at the points $(s_1, 0, 0)$, $(s_2, 0, 0), (0, s_3, 0), (0, s_4, 0), (0, 0, s_5)$ and $(0, 0, s_6)$.

3. FORMULATION

Let us begin by noting that the probability density functions of x^j can *be written* in a product form, as

$$p(x^j) = \prod_{i=1}^{3} \phi(x_i^j; \mu_{x_i^j}, \sigma_{x_i^j}^2) \tag{3.1}$$

where

$$j = 1, 2, 3..$$

The probability that the first round hits the target T, $P(F)$, is then given by

$$P(F) = \int_T p(x^1) dx^1 \tag{3.2}$$

where the integral symbol denotes the triple integral over T.
Substituting (3.1) into (3.2), we obtain $P(F)$ expressed in terms of error functions as follows:

$$P(F) = \prod_{i=1}^{3} \left[\text{erf}(\alpha_{2i}) - \text{erf}(\alpha_{2i-1}) \right] / 2 \tag{3.4}$$

where

$$\alpha_{2i-1} = c_0 (s_{2i-1} - \mu_{x_i^1}) / \sigma_{x_i^1}$$

$$\alpha_{2i} = c_0 (s_{2i} - \mu_{x_i^1}) / \sigma_{x_i^1}.$$

for $i = 1, 2, 3$ and $c_0 = 1/\sqrt{2}$

Let x^1 designate the coordinates of the first round of a burst, and x^2 denote that of the subsequent rounds. The coordinates for the offset of the mean center of impact of the subsequent rounds, which is $\mu_{x^2} = (\mu_{x_1^2}, \mu_{x_2^2}, \mu_{x_3^2})$, from the first round is denoted by $x^3 = (x_1^3, x_2^3, x_3^3)$, where $x_i^3 = \mu_{x_i^2} - x_i^1$, $i = 1,2,3$ or simply $x^3 = \mu_{x^2} - x^1$. $\qquad$ (2.1)

By the assumptions stated in the previous section, the density functions of the three mutually independent, normally distributed random variables, X_i, $i = 1,2,3$, can be written as $\phi(x_i; \mu_{x_i}, \sigma_{x_i}^2)$, where μ_{x_i} and $\sigma_{x_i}^2$ are the mean and variance, respectively, of the distribution corresponding to the random variables. It follows that $\mu_{x_i^1}$, $i = 1,2,3$, can be considered as the actual aim point.

2.2 The Problem

The foregoing considerations enable us to state the first part of the problem as follows: given a target T, defined by the relation

$$T = \{x \varepsilon E^3 \mid s_1 \leq x_1 \leq s_2, \ s_3 \leq x_2 \leq s_4, \ s_5 \leq x_3 \leq s_6\} \qquad (2.2)$$

and the input parameters— $\mu_{x_i^j}$, $i = 1,2,3$ $j = 1,3$; $\quad \sigma_{x_i^j}^2$, $i,j = 1,2,3$ N, the total number of rounds for the burst; and $P(I|H)$, the probability of incapacitation per hit, — find $E(H)$ and $P(I)$.

3. FORMULATION

Let us begin by noting that the probability density functions of x^j can be written in a product form, as

$$p(x^j) = \prod_{i=1}^{3} \phi(x_i^j; \mu_{x_i^j}, \sigma_{x_i^j}^2) \qquad (3.1)$$

where

$$j = 1,2,3..$$

The probability that the first round hits the target T, $P(F)$, is then given by

$$P(F) = \int_T p(x^1) dx^1 \qquad (3.2)$$

where the integral symbol denotes the triple integral over T. Substituting (3.1) into (3.2), we obtain $P(F)$ expressed in terms of error functions as follows:

$$P(F) = \prod_{i=1}^{3} \left[\mathrm{erf}(\alpha_{2i}) - \mathrm{erf}(\alpha_{2i-1}) \right]/2 \qquad (3.4)$$

where

$$\alpha_{2i-1} = c_0 (s_{2i-1} - \mu_{x_i^1})/\sigma_{x_i^1}$$

$$\alpha_{2i} = c_0 (s_{2i} - \mu_{x_i^1})/\sigma_{x_i^1} .$$

for $i = 1,2,3$ and $c_0 = 1/\sqrt{2}$

For a given μ_{x2} the probability that a subsequent round hits T, denoted by $P(S|\mu_{x2})$, is

$$P(S|\mu_{x2}) = \int_T p(x^2)dx^2 \qquad (3.5)$$

The probability that a subsequent round hits T given x^1, denoted by $P(S|x^1)$, is

$$P(S|x^1) = \int_{E^3} p(x^3) \ P(S|\mu_{x2})d\mu_{x2} \qquad (3.6)$$

where $d\mu_{x2} = d\mu_{x_1^2} \cdot d\mu_{x_2^2} \cdot d\mu_{x_3^2}$ and the integral symbol denotes the triple integral over E^3. Using (2.1) and (3.5) in (3.6), we can rewrite (3.6) as follows:

$$P(S|x^1) = \int_T \int_{E^3} p(\mu_{x2}-x^1) \ p(x^2)d\mu_{x2} \ dx^2 \qquad (3.7)$$

The probability that a subsequent round hits T, denoted by $P(S)$, is then

$$P(S) = \int_{E^3} p(x^1) \quad P(S|x^1)dx^1 \qquad (3.8)$$

Substitution of (3.7) into (3.8) yields the following:

$$P(S) = \prod_{i=1}^{3} (\text{erf}(\beta_{2i}) - \text{erf}(\beta_{2i-1}))/2 \qquad (3.9)$$

where

$$\beta_{2i-1} = b(s_{2i-1} - \mu_{x_i^3} - \mu_{x_i^1})/\bar{\sigma}_{x_i}$$

$$\beta_{2i} = b(s_{2i} - \mu_{x_i^3} - \mu_{x_i^1})/\bar{\sigma}_{x_i}$$

for $i = 1,2,3$ and $\bar{\sigma}_{x_i} = (\sigma_{x_i^1}^2 + \sigma_{x_i^2}^2 + \sigma_{x_i^3}^2)^{1/2}$

The expression for $E(H)$ can now be readily obtained for a burst of N rounds by substituting equations (3.4) and (3.9) in the following relation:

$$E(H) = \sum_{i \in \Omega} X_\Omega \, P(H_i) \tag{3.10}$$

where $P(H_i)$ is the probability that the ith round in a burst of N rounds hits T and the "characteristic function" X on the index set $\Omega = (1,2,3, \cdots ,N)$ is defined by the formula

$$X_\Omega = \begin{cases} 1 & \text{if} \quad i \in \Omega \\ 0 & \text{if} \quad i \notin \Omega \end{cases}$$

It follows that

$$E(H) = P(F) + (N-1) \, P(S) \tag{3.11}$$

To seek $P(I)$ we proceed in the following manner. It is wellknown [5] that for any two measurable events

$$P(\overline{I}_f \, \overline{I}_s) + P(I_f \, \overline{I}_s) = P(\overline{I}_s) \tag{3.12}$$

where

$P(\overline{I}_f \, \overline{I}_s)$ denotes the probability that the first projectile and the subsequent projectiles do not incapacitate the target T.

$P(I_f \overline{I}_s)$ denotes the negation of $P(\overline{I}_f \overline{I}_s)$ on the first event.

$P(\overline{I}_s)$ denotes the probability that none of the subsequent projectiles incapacitates T.

Hence,

$$P(\overline{I}) = P(\overline{I}_s) - P(I_f \, \overline{I}_s) \tag{3.13}$$

For a given μ_{x2}, the probability that none of the $N-1$ subsequent projectiles incapacitates T, denoted by $P(\bar{I}_s|\mu_{x2})$, is given by

$$P(\bar{I}_s|\mu_{x2}) = \prod_{i=2}^{N} [1 - P(I|H_i)P(H_i|\mu_{x2})] \qquad (3.14)$$

The probability $P(\bar{I}_s|x^1)$ that none of the subsequent projectiles incapacitates T given x^1 is given by

$$P(\bar{I}_s|x^1) = \int_E P(x^3)P(\bar{I}_s|\mu_{x2})d\mu_{x2} \qquad (3.15)$$

It follows that

$$P(\bar{I}_s) = \sum_{k=0}^{N-1} \binom{N-1}{k} [-P(I|H)]^k \prod_{i=1}^{3} \int_E \left\{ \frac{e^{-\frac{1}{2}F_4}}{\sqrt{2\pi}\bar{\sigma}_{x_i}^{2^k}} \right.$$

$$\left. \times [erf(\beta_{2i}) - erf(\beta_{2i-1})]^k \right\} d\mu_{x_i^2} \qquad (3.16)$$

where

$$F_4 = (\mu_{x_i^1} - \mu_{x_i^2} + \mu_{x_i^3})^2/\bar{\sigma}_i$$

$$\bar{\sigma}_{x_i} = \sigma_{x_i^1}^2 + \sigma_{x_i^3}^2$$

$$\beta_{2i-1} = (s_{2i-1} - \mu_{x_i^2})/\sqrt{2}\,\sigma_{x_i^2}$$

$$\beta_{2i} = (s_{2i} - \mu_{x_i^2})/\sqrt{2}\,\sigma_{x_i^2}$$

$$i = 1,2,3.$$

Here we have used the assumption that for a given hit the probability of incapacitation is equally probable; i.e. $P(I|H) = P(I|H_i)$, for all $i \in G$

To evaluate $P(I_f \overline{I}_s)$, we note that

$$P(I_f \overline{I}_s) = P(I|H)P(F \overline{I}_s) \tag{3.17}$$

where

$$P(F \overline{I}_s) = \int_T P(x^1)P(\overline{I}_s|x^1)dx^1 \tag{3.18}$$

After some computations on the last integral, we arrive at

$$P(I_f \overline{I}_s) = P(I|H) \sum_{k=0}^{N-1} \binom{N-1}{k} [-P(I|H)]^k \prod_{i=1}^{3}$$

$$\int_E \left\{ \frac{e^{-\frac{1}{2}F_e}}{(2^{k+1})\sqrt{2\pi}\,\sigma_{x_i^1}} \, [erf(\beta_{2i}) - erf(\beta_{2i-1})]^k \right. \tag{3.19}$$

$$\left. \cdot [erf(a\,\overline{s}_{2i}) - erf(a\,\overline{s}_{2i-1})] \right\} d\mu_{x_i^2}$$

where

$$a = \sqrt{\overline{\sigma}_{x_i}}/\sqrt{2}\,\sigma_{x_i^1}\,\sigma_{x_i^2}$$

$$\overline{s}_{2i-1} = [s_{2i-1} - \frac{\sigma_{x_i^3}^2 \mu_{x_i^1} + \sigma_{x_i^1}^2(\mu_{x_i^2} - \mu_{x_i^3})}{\overline{\sigma}_{x_i}}]$$

$$\overline{s}_{2i} = [s_{2i} - \frac{\sigma_{x_i^3}^2 \mu_{x_i^1} + \sigma_{x_i^1}^2(\mu_{x_i^2} - \mu_{x_i^3})}{\overline{\sigma}_{x_i}}$$

$$i = 1,2,3.$$

$P(I)$ can be obtained once $P(\overline{I})$ in (3.13) is determined; ie.

$$P(I) = 1 - P(\overline{I}) \tag{3.20}$$

4. NUMERICAL EXAMPLES

For illustration, a regular parallelepiped target of 6x12x15 inches is considered first. As a special interest we assume that the target is zeroed in at the center of its front surface base line i.e. $-3 \leq x_1 \leq 3$, $0 \leq x_2 \leq 12$, and $0 \leq x_3 \leq 15$. A rectangular target of dimensions $-3 \leq x_1 \leq 3$ and $0 \leq x_3 \leq 15$ is also considered as a special case. It provides a base for verification with that given in [3].

Table 1 gives the values of the input parameters- N, $\sigma_{x_i^j}$ and $\mu_{x_i^j}$ - for three different cases. Table 2 shows the results of the effectiveness parameters for the two targets. The value of $P(I|H)$ is taken as 0.85 for finding the values of $P(I)$.

Numerical integrations were computed by using Simpson's rule. For each integration, equally spaced increments are used over $6\sigma_{x_i^j}$ ranges. Any numerator value less than 10^{-16}, as well as any value for the difference of two error functions less than 10^{-6}, is taken to be zero.

5. DISCUSSIONS & CONCLUSIONS

An effectiveness model for machine gun bursts on volume targets has been developed with three-trivariate normal distributions in this report. It is noted that for targets having special configurations, the corresponding models can be readily obtained by an appropriate coordinate transformation[7] An area target model is obtainable as a special case.

For instance, the model so obtained for regular parallelepiped targets is an obvious extension of the three-bivariate model rendered by [3]. This fact also lead us to note that $P(F)$ and $P(S)$ for a volume target can at most assume only one-half of the corresponding values for an area target.

The conventional notion of an impact point becomes rather nebulous when carried over to a volume target. Further pursuance for a rigorousness development of this notion may be necessary. The reader is referred to [6] for such an effort.

Since the essential premises used for this modeling in extending the random phenomena from two dimensions into the third dimension is a strong one. Users of the model should comprise with the stipulations.

TABLE 1

Input Parameters: No. of rounds, means, and standard deviations

Case	N	j	$\sigma_{x_1}^j$	$\sigma_{x_2}^{j*}$	$\sigma_{x_3}^j$	$\mu_{x_1}^j$	$\mu_{x_2}^{j*}$	$\mu_{x_3}^j$
1	6	1	0.001	0.001	0.001	0.000	0.000	0.000
		2	3.903	4.250	4.671	-3.224	-5.507	-2.790
		3	5.839	9.800	13.27			
2	15	1	0.001	0.001	0.001	0.000	0.000	0.000
		2	3.774	4.0175	4.261	-3.699	-0.634	2.431
		3	4.515	6.059	7.601			
3	15	1	0.001	0.001	0.001	0.000	0.000	0.000
		2	3.057	6.055	9.053	-0.118	0.3285	0.775
		3	3.385	5.6535	7.922			

* Zero for area target
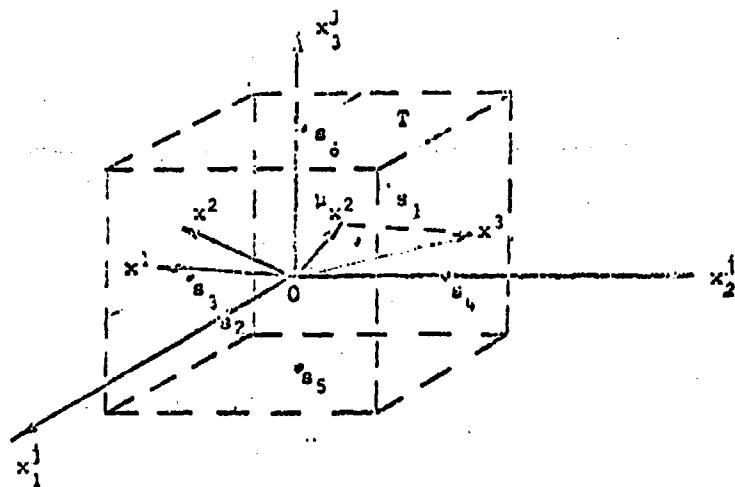
Fig. 1 Coordinates of x^j, $j = 1,2,3$, and a
Solid Parallelepiped Target T

Table 2

Effectiveness Parameters: Solid
Parallelepiped and Rectangular Targets

Target	Solid Parallelepiped			Rectangular		
Parameter/case	1	2	3	1	2	3
$P(F)$	0.250	0.250	0.250	0.500	0.500	0.500
$P(S)$	0.015	0.079	0.095	0.061	0.186	0.219
$E(H)$	0.326	1.356	1.592	0.805	3.103	3.578
$P(1^+)$	0.753	0.670	0.715	0.828	0.881	0.927
$P(I)^*$	0.736	0.643	0.668	0.793	0.850	0.901

REFERENCES

1. "Table of Salvo Kill Probabilities for Squad Targets", U.S. Department of Commerce National Bureau of Standards, Applied Mathematics Series 44, December, 1954.

2. F. A. Malinoski, "A Summary of Mathematical Models in Hit and Incapacitation Probability Analysis of Small Arms Weapon Systems", Report R-1831, U. S. Army Frankford Arsenal, December,1966, AD 379783 (C).

3. H. K. Fallin, "Analysis of Machine Gun Burst Dispersion Data with Corresponding Effectiveness Model", AMSAA, TM-33, Aberdeen Proving Ground, Maryland, July 1969.

4. H. K. Fallin, "A Mathematical Model for Machine Gun Effectiveness", BRL TN1622, August 1966.

5. E. Purzen, "Modern Probability Theory and its Applications", John Wiley & Sons, Inc., New York, 1962.

6. H.M. Hung and J.T.Wong, "On the Ballistics Effectiveness as Characterized by Maximal Trajectory Information", Systems Analysis Directorate, US Army Weapons Command, Rock Island, Illinois.

7. T. H. M. Hung and J. T. Wong, "A Machine Gun Burst Effectiveness Model for Volume Target", Systems Research Division, RD&E Directorate, US Army Weapons Command, Rock Island, Illinois.

SIMULATION OF SUBSURFACE NUCLEAR EXPLOSIONS
WITH CHEMICAL EXPLOSIVES

Donald E. Burton, and Edward J. Leahy
Lawrence Livermore Laboratory, Livermore, California

ABSTRACT

A series of field experiments are being designed and theoretical
studies are being conducted to select a chemical explosive and to develop
explosive charge configurations and synthetic fallout material to simu-
late subsurface, sub-kiloton nuclear cratering explosions. The studies
and experiments are directed toward determining the effects, the type
and degree of stemming (full stemming, water stemming, and no stem-
ming) have on the size of nuclear craters, the vented radioactivity, and
the extent of the resulting fallout pattern. Such information for detona-
tions in a variety of geologic media is required if nuclear explosives are
to be developed as a civil and military construction tool.
The problems associated with simulating these effects through the
use of chemical explosives are discussed, and partial solutions are pre-
sented. The theoretical and experimental programs will study the relative
effects of the different stemming configurations in the chemical explosive
case and relate the results to the nuclear case. It is suggested that
these investigations will lead to optimum design criteria for simulation
experiments and to a means of inferring the information which cannot be
simulated.

1. INTRODUCTION

The development of nuclear explosives as a tool for civil and military construction requires testing in a variety of geologic media and near-surface emplacement configurations. Chemical explosives are a convenient and economical way to model nuclear explosive effects. The U. S. Army Engineer Explosive Excavation Research Office (EERO), Livermore, California, a part of the U. S. Army's Waterways Experiment Station at Vicksburg, Mississippi, and the Lawrence Livermore Laboratory (L^3), Livermore, California, are designing a series of field experiments to be used in conjunction with numerical modeling calculations to develop such a simulation technique. Related numerical calculations are also being conducted by Systems, Science, and Software (S^3), La Jolla, California. The combined effort is called Project DIAMOND ORE.*

In the development of the technology to permit employment of nuclear explosives as a tool for use in both civil and military construction, the influence of burial depth and stemming material on phenomena associated with nuclear cratering must be determined. The construction of engineering structures such as railroad cuts, storage areas, and obstacle craters requires a detailed knowledge of ground motion, limits of material failure, and final crater size. Safety considerations require knowledge of airblast overpressures, seismic motions, missile throwout ranges, and radioactive fallout distribution. The DIAMOND ORE program is therefore directed toward modeling as many of these effects as possible with the primary emphasis placed on crater size and radioactivity modeling.*

---

* Project DIAMOND ORE IS JOINTLY FUNDED BY THE Defense Nuclear Agency and the Office of the Chief of Engineers, US Army.

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

-128-

II. SIMILITUDE AND CRATERING PHENOMENOLOGY

The key to the successful simulation of nuclear cratering events through the use of high explosives (HE) is the understanding of the cratering phenomenology which results from both energy sources. Although the quantitative aspects of simulation will not be completely known until the termination of Project DIAMOND ORE, the following qualitative aspects supply a starting point for formulating simulation criteria.

(1) Early-time Energy Source Behavior

All of the energy from a nuclear device is released in less than a microsecond, vaporizing the device materials and its canister. (See Point 1 in Figure 1). As this high-pressure, high-temperature gas pushes on the walls of the emplacement cavity, a strong shock wave propagates through the surrounding rock, depositing sufficient energy to cause vaporization. Butkovich[1] has shown that, in silicate rock, about 70 tons of rock are vaporized per kiloton of yield (1 kt = 10^{12} cal). In the region of vaporization, the response of the rock to the shock is such that a lower limit of about 20% of the device energy is lost in internal heating and will not be available to do mechanical work at late times. As the shock wave propagates outward, it continues to melt the rock and shock-vaporize water contained in pores and cracks. Shock vaporization of silicate rock and water occurs at pressures of the order of a megabar and 100 kilobars, respectively. The spherically diverging shock wave and the expanding mass of vaporized rock and water are the principal agents in the formation of nuclear
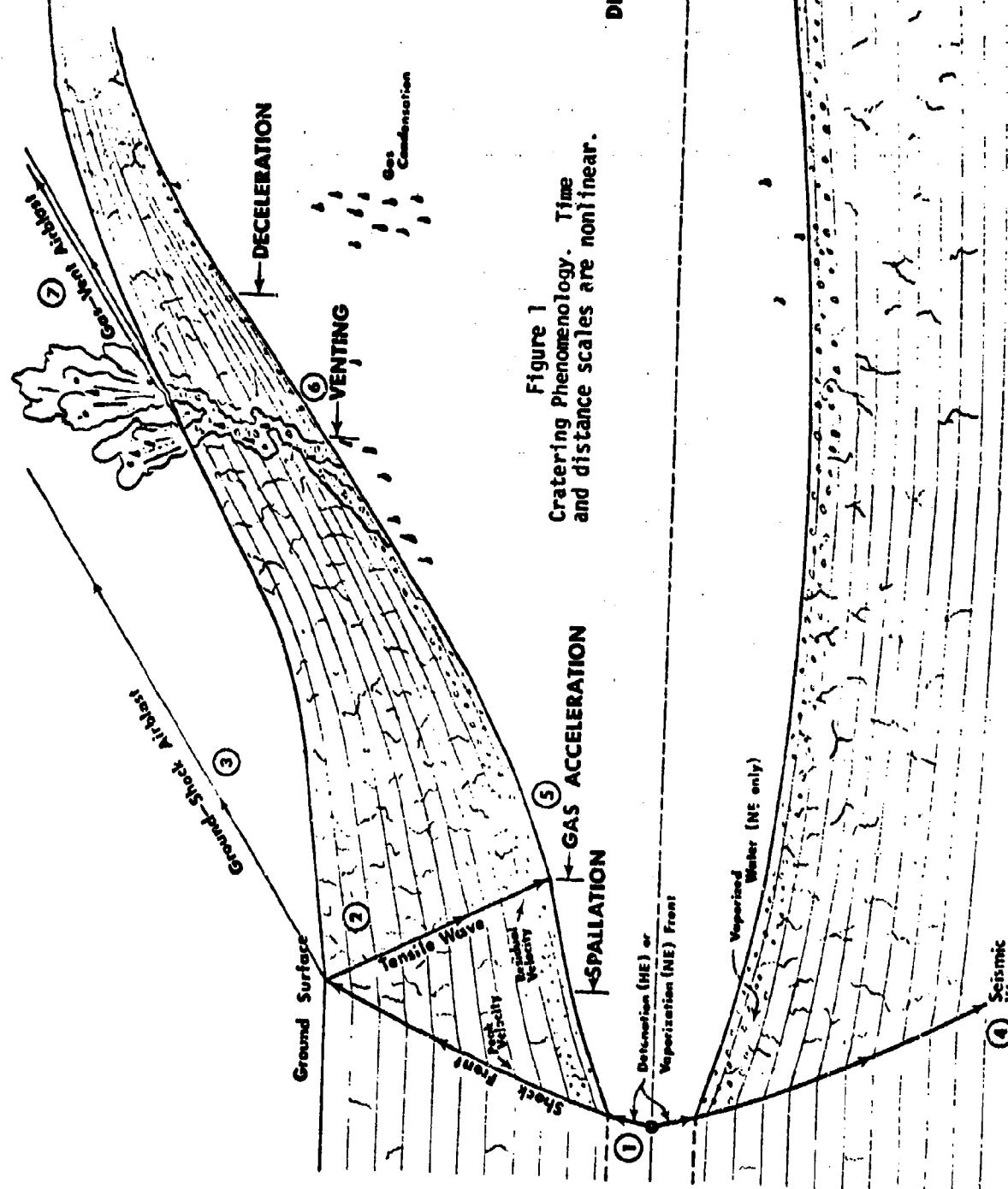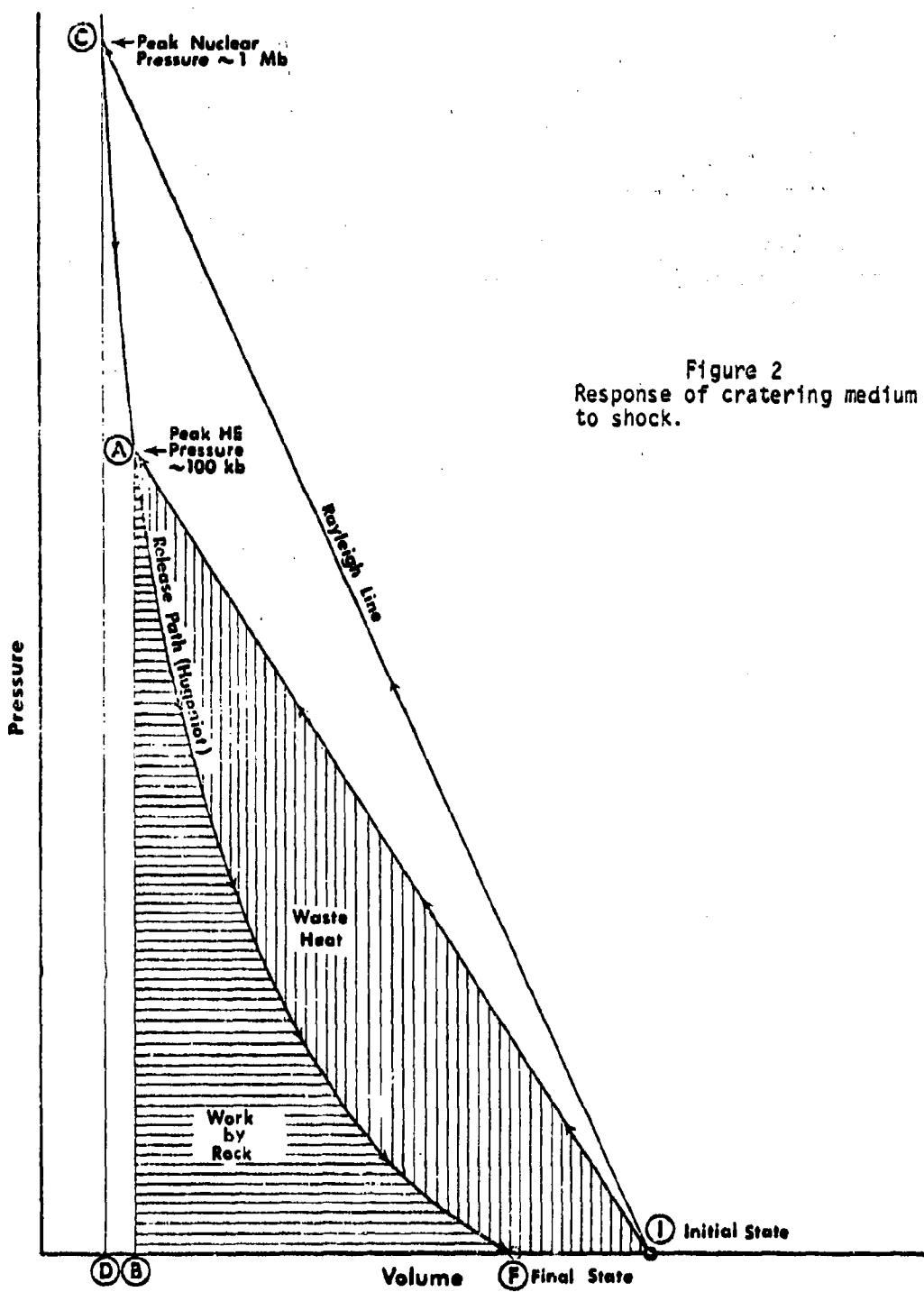
Figure 1
Cratering Phenomenology. Time and distance scales are nonlinear.

craters.

In the case of chemical explosives, the detonation is completed in a time of the order of a millisecond, the exact value depending upon the yield and type of explosive. Because the average pressures within the cavity formed by the gaseous detonation products are typically less than 100 kilobars, effectively no water or rock is shock vaporized in the medium surrounding the HE cavity.

The HE detonation products play a role in HE cratering which is analogous to that of the vaporized rock in nuclear cratering. In a modeling experiment, complete similitude cannot be achieved if the pressure-volume adiabat for the chemical explosive detonation products does not coincide with that of the nuclear cavity gas. This condition obviously cannot be met because the nuclear gas does work on its surroundings in the pressure range between 100 kb and 1 Mb, while the HE does not. Because of the large initial pressure difference, the coincidence of the HE and vaporized rock adiabats in the high pressure regions can no longer be used as a similitude requirement.

However, successful simulation of the cratering mechanism hinges upon transferring not the same total energy to the medium as the above argument would require, but rather the same kinetic energy. The peak pressure and response of the cratering medium to the shock wave governs the amount of energy lost in shock heating and, consequently, the amount of kinetic energy transferred to the medium. The response of a given material to a shock is characterized by a release path or Hugoniot in pressure-volume space (see Figure 2). In an HE detonation, a region of material near the HE cavity is shocked from point I to A, receiving energy

Figure 2
Response of cratering medium
to shock.

represented by the area IAB. As the pressure decreases, the material unloads along the Hugoniot, doing work ABF on the surroundings and retaining energy AFI as internal energy or "waste heat". In a nuclear detonation, the medium is subjected to a much greater pressure and receives greater energy ICD. However, most of this energy is lost as waste heat ICF. Consequently, although the nuclear case involves much greater pressures than the HE case, only a relatively small amount of additional energy (BCD) is made available for the cratering process.

In other words, the very high pressure shock wave resulting from a nuclear detonation does not couple efficiently with the medium, giving rise to large energy losses through shock heating. (The presence of walls allows a small but significant quantity of this heating energy to eventually return to the system in a mechanically usable form). As the shock wave propagates outward, this energy loss results in the rapid attenuation of the shock peak to levels attainable with high explosives. Because of the lower loss rate of the HE shock wave, cratering similitude can then be achieved even though the adiabats for the vaporized rock and the HE detonation products are not identical in the high-pressure region. Furthermore, because of the difference in total energy loss between the two cases, the HE charge can be selected to have a much lower energy yield than the nuclear device. However, it must be understood that effects which depend upon the material temperature and effects occurring in the region very near the cavity cannot be modeled.

(2) Spallation

In both the nuclear and HE cases, as the shock wave intersects material layers and finally the surface, it is reflected back towards the

-133-

cavity in the form of rarefactions or tensile waves. As the tensile wave passes downward through the medium, the material,broken by compression, tension, or shear, spalls off in layers which then execute an approximate ballistic trajectory. (See point 2 in Figure 1).

The depth from which material is ejected by spall is a function of the material strength and kinetic energy transferred to the mound by the shock wave. In simulating the spall phenomena, primary consideration must be given to matching not the peak particle velocities which are short-lived (except at the surface), but rather the residual particle velocities (those particle velocities which remain after the shock peak has subsided) throughout the mound. This criterion is approximately equivalent to matching the total mound kinetic energies produced by the shock front. Preliminary calculations suggest that only 15 to 20% of the total nuclear device yield is converted to mound kinetic energy prior to the gas acceleration phase, whereas this figure may be as high as 30 to 40% for an HE such as nitromethane.

(3) Ground-Shock-Induced Airblast

When the shock hits the free surface, the top surface layers will spall off at approximately twice the peak velocity of the particles at the same distance from shot point but not at the surface. (Point 3 in Figure 1). The overall motion of the ground surface acts as a piston and couples with the atmosphere, producing "ground-shock-induced" airblast. The simulation of this effect then requires that the HE and nuclear explosions produce the same peak particle velocities at the ground surface. This criterion conceivably conflicts with the previously stated kinetic energy criterion. Although it may be possible to match both, such circumstances

depend upon the rapid attenuation of the sharp peak of the nuclear shock wave relative to the peak of the HE shock wave. Whether sufficient attenuation has occurred by the time the wave strikes the surface depends upon the material properties and the shot depth.

(4) Seismic Effects

As the shock wave (and associated particle motion) propagates outward in the horizontal direction, it attenuates rapidly with distance because of the spherical divergence and because of energy deposition in the medium. Dispersive effects round off the sharp peak of the shock front through the attenuation of high frequencies; and, as the shock amplitude decreases, the material begins to behave elastically. The disturbance continues to move outward as a seismic pulse traveling at acoustic velocities. Because of the dispersion, the amplitude of the particle motion is more directly related to the residual particle velocities than to the peak particle velocities. Because of the dispersion, the peak and residual particle velocities will behave in a similar manner at large distances. For this reason, it is anticipated that modeling of seismic motion will be achieved by a similitude based upon the kinetic energy criterion.

(5) Gas Acceleration

As the rarefaction impinges on the cavity, the stresses in the rock at the roof of the cavity are relieved; and the cavity begins an accelerated asymmetrical growth toward the surface, producing the so-called gas acceleration phase of the cratering process. (Point 5 in Figure 1.) The magnitude of this effect depends primarily upon the time at which the rarefaction intersects the cavity (and therefore the shot depth), the cavity pressure at this time, and the expansion of the cavity gas below this pressure

(i.e., the low-pressure gas adiabat). As the cavity expands, the pressure and temperature decrease. Any drop in temperature below about 2000° K will be accompanied by the partial condensation of the vaporized rock gas, bringing about an accelerated drop in pressure. Such an effect will not be experienced by the essentially noncondensable HE detonation products. Preliminary calculations indicate that for the same yield the pressure and available energy in a nitromethane cavity gas at relatively late times may be several times that of the nuclear cavity bases. For nuclear detonations in wet materials, the steam produced by shock vaporization will behave as a noncondensable gas and will give rise to somewhat higher late-time pressures than in comparable dry materials.

As indicated earlier, the simulation of spallation requires the matching of the mound kinetic energies rather than the high-pressure gas adiabats. The simulation of the gas acceleration phase, however, requires the coincidence of the adiabats at the low pressure end. While these are not necessarily mutually exclusive criteria, they may not be simultaneously realizable with present-day explosives. The modeling of either spallation cratering or gas-acceleration cratering presents no conceptual problems. However, the modeling of nuclear events in which both processes are significant may present many difficulties. Certainly, it is always possible to select an HE yield which will produce a crater of the same volume as that produced by a spallation/gas-acceleration nuclear shot. But, some of the secondary effects will not be properly modeled; the ratio of radius to depth, for example, may be in error. The magnitude of these errors can only be found through HE experiments and extensive numerical modeling.

(h) Depth of Burst and Relative Crater Sizes

In very shallow shots, *the* overburden spalls off quickly so that the
cavity gases vent and do work on the atmosphere rather than the mound.
In somewhat deeper shots, the early spallation is reinforced by the later
gas acceleration mechanism. As the depth of burst is increased still
further, the shock wave becomes weaker at the surface, causing less
spallation; and gas acceleration begins to play the dominant role. Although
it is commonly believed that maximum crater size is obtained for shot
depths at which spallation and gas acceleration contribute equally, this
has not been firmly established, and only experiment or precise numerical
modeling can answer this question for a given medium.

Thus, the spallation and gas acceleration mechanisms combine to
produce a variation in crater dimensions with depth for a particular
energy yield and material. Generally, the crater radius and depth peak at
approximately the same burial depth which is termed the "optimum depth of
burst". Because of the different emphasis placed on spallation and gas
acceleration by the different energy sources, the "cratering curves" for
a high explosive and a nuclear explosive (or, for that matter, other high
explosives) cannot be expected to have the same shape, and therefore the
same optimum depth of burst. As has been indicated several times, the
best approximation should occur when the same kinetic energy is imparted
to the mound by the shock wave and when the cavity gases have the same
late-time pressure. If this criterion is met, the cratering curves should
coincide, providing the HE and nuclear yields are redefined to take into
account the energy loss due to shock heating of the medium. This redefini-
tion is, of course, media-dependent to some extent. If the condition cannot

-137-

be met, one must resort to simulating a fixed-yield nuclear explosive at different depths with HE charges of varying yields which will depend upon the predominant cratering mechanism at that depth. Use of this approach will result in the imperfect simulation of some of the secondary nuclear effects.

(7) Missile Hazards

As the maximum missile ranges are usually produced by the larger blocks of material which have been spalled off the surface by the shock wave, simulation involves the same criteria and consequences as discussed for ground-shock-induced airblast. An exception to this may occur in soils for which the top layers are significantly weaker than the deeper materials. Under these conditions, the more energetic missiles may originate from the deeper layers.

(8) Gas-Vent-Induced Airblast

The effects of the cavity gas extend beyond providing a boost to the mound. If the mound breaks up before the cavity pressure has dropped to one atmosphere, the cavity gases do work on the air to produce gas-vent airblast. (Point 7 in Figure 1). Simulation of this effect involves reproducing the nuclear cavity pressure and energy at vent time and, consequently, matching the adiabats in the low-pressure range. This presents many problems, particularly in shallow shots in which spall is the primary cratering mechanism. That is to say, matching the adiabats to reproduce gas-vent airblast may give rise to a mismatch in the spall-producing portion of the adiabats possibly to the extent of affecting the vent time.

(9) Venting[2]

In nuclear events, the behavior of the cavity gas is intimately related to the venting of radioactive fission products and induced radionuclides which later appear as fallout. As the mound grows and begins to dissociate, those radioactive cavity gases which are still uncondensed filter or are injected through fissures in the mound. The violence with which the gas is released to the atmosphere will depend upon the cratering mechanism at the time of release: (a) if gas acceleration is the predominant mechanism, then a large cavity pressure at late times can be expected to give rise to a relatively large vent; (b) in a shallow, spall-produced crater, the spall depth can extend into the cavity region. Under these circumstances, the venting should be large and may be relatively independent of cavity pressure. On the other hand, if spall is the primary cratering agent and the spall depth does not extend into the cavity region, then a relatively small vent is to be expected. Cavity gas will still be forced through fissures by the late-time cavity pressure, but to a lesser extent than in the previous cases. Clearly, the _complete_ simulation of the venting mechanism requires comparable late-time cavity pressures and therefore similar adiabats for the HE and vaporized rock in the lower pressure region, although it is possible that this condition may be overly restrictive in that the venting mechanisms may not be sensitive to the exact cavity pressures involved.

The cavity pressure at vent time can also have a significant effect on the composition of the fallout. This arises because the radioactive fission products and induced activities entrained in the cavity gas behave in chemically and physically varied manners. The "refractory"

radionuclides belong to mass chains which, on the average, form oxides with relatively high condensation temperatures between about 1100 and 1700°C[3]. If the cavity tormperature drops below this range prior to venting (corresponding to a lower cavity pressure and less violent vent), the refractory radionuclides will condense along with the vaporized rock to form volume-distributed droplets which are preferentially surface-deposited on the large particles of the dissociating mound. Radionuclides belonging to the "volatile" mass chains will not condense until much lower temperatures are reached. Presumably, the condensation of volatiles will occur at a late time when there is a preponderence of fine particles, so that the volatiles will be preferentially deposited on smaller particles than the refractories. This difference in behavior is termed "fractionation", and gives rise to a distribution of radionuclides in the cloud and fallout which differs markedly from that distribution produced by the nuclear source.

If, on the other hand, the cavity pressure is high at the time of vent, neither refractories nor volatiles will be afforded the opportunity to condense prior to being injected through the mound into the atmosphere. Under these conditions, there will be little preferential condensation on particles of any given size, and no marked fractionation will occur. (Notable exceptions to this are radionuclides which behave predominantly as noncondensable gases.)

Consequently, it is legitimate to speak of a "vent fraction" (or fraction of produced radionuclides which are vented into the cloud and fallout) in the case of violent venting. However, in the case of a low-pressure vent, one must speak of a vent fraction for each individual mass chain or, at best, a separate vent fraction for refractory, volatile, and

intermediate classes of radionuclides.

As it is obviously impractical to attempt simulation of a nuclear fallout pattern with high explosives, the modeling must be confined to reproducing a parameter such as the vent fraction(s) which will allow an equivalent fallout pattern to be generated by numerical techniques. In the case of high-pressure venting, this simulation can be accomplished in a relatively straightforward manner. In principle, one need only emplace some sort of traceable material uniformly throughout the high explosive and then recover it from the region outside the crater. In practice, there are many problems which must be overcome; these will be described in more detail later. Again, there is the everpresent restriction that the HE detonation products and the nuclear cavity gases must have approximately the same late-time pressure. Simulation of low-pressure venting necessarily involves the use of different types of tracer materials. It is doubtful that particulate tracers will satisfactorily simulate the volatiles under these conditions. Ideally, one needs tracers which will go through the same physical phases at the same times as the radionuclides but at the much lower HE temperatures. At the present time, considerations of this nature appear to be beyond the state-of-the-art.

(10) Cloud Formation

Cratering events normally exhibit a cloud structure comprising of a main cloud and a base surge. This structure is intimately related to the distribution of fallout produced by nuclear events. With the complete dissociation of the mound, the material is ejected along ballistic trajectories. The primary mechanism for the formation of the base surge is thought to be the potential energy in the suspended aerosol formed from

-141-

the lofted mound material. The formation is interpreted physically as a gravity flow of the aerosol, during which energy is transformed from potential energy to kinetic energy of the aerosol and then to work done in displacing environmental air[2]. Correct simulation of the base surge then follows from the correct simulation of the late-time mound kinetic energy whether caused by spallation or gas acceleration.

The formation of the main cloud is keyed to the cavity pressure at vent time; and, the absence of this cloud can be expected if a violent vent does not occur. With the violent release of the cavity gases, the main cloud rises from ground zero and grows primarily in the vertical direction, entraining fine particulates from the mound. The growth is influenced by the initial momentum of the hot cavity gas bubble, buoyant forces on the bubble, the internal friction produced by turbulence, the external friction and pressure exerted on the bubble by the environment, and the entrainment of mass in the form of air and particulates.

If the main cloud were formed by an adiabatic process, then the matching of the cavity pressures and adiabats would be the key factor in arriving at an HE simulation. However, in reaching stabilization, the cloud exchanges energy with the environmental air. As the HE cannot attain the temperatures of the nuclear cavity gas, it is doubtful that there exists a main-cloud simulation criterion which is compatible with the previously stated cratering simulation criteria.

(11) Effects of Stemming Materials

(a) Solid Stemming. Insofar as the material with which the emplacement hole is stemmed is a reasonable approximation to the in-situ material, the stem will not have a profound effect upon the early spallation stages of

crater dovolopment. However, as the cavity expands, the original borehole also enlarges, first near the charge and later near the ground surface. Solid stemming may tend to lose contact with the wall surfaces, thereby permitting premature venting of the cavity gases with subsequent loss of energy available for the gas-acceleration stage. This effect can of course be reduced by special emplacement designs. If other simulation criteria are met, it is likely that solid-stemming materials will behave comparably in the HE and nuclear cases.

(b) Liquid Stemming. The use of water stemming has been tested using high explosives with the possible result of increased cratering efficiency. This has been attributed to the ability of liquid stemming to flow laterally, maintaining contact with the walls of the borehole[4]. This concept cannot be simply extended to the nuclear case because different physical principles come into play, and one can only speculate on the consequences. As indicated earlier, the amplitude of the nuclear shock wave is sufficient to shock-vaporize water considerably beyond the limit of rock vaporization. This will cause the formation of a steam bubble in the lower portion of the stem which might contribute to the premature ejection of the comparatively low-density water stem with the concommitant early venting of the radioactive cavity gases. Although the design of an HE configuration to simulate this effect must be based upon numerical calculations, one can speculate that this could be accomplished by prechambering part of the HE in the bottom part of the water stem.

(c) Air Stemming. There is limited evidence to indicate that the absence of stemming in HE detonations does give rise to a degradation in cratering efficiency[4, 5]. Once initiated by the HE or nuclear detonation, the shock

propagates radially outward not feeling the presence of the unstemmed hole except in the direction along the axis of the hole. For this reason, the spallation mechanism set into motion by the shock wave should be relatively unaffected by the lack of stemming. However, all effects which depend upon the cavity gas may be greatly affected.

In the simplest approximation, the borehole can be considered to be a high-pressure shock tube. The application of ideal shock tube theory[6] to the nuclear case indicates that the initial velocity of the cavity-gas/air interface in the borehole may be greater than 20 meters per millisecond. (During the Marvel experiment, a velocity of about 100 m/ms was observed[7].) Similar calculations indicate that a nitromethane high explosive may exhibit a velocity of approximately 10 m/ms. At these speeds, the interface will outdistance the shock wave in the adjacent material, originating a secondary shock wave from the interior of the borehole. Should such interface velocities be sustained, then a significant mass of the cavity gas will be ejected up the borehole, the rate depending upon the hole diameter. However, such rates cannot be sustained because a number of effects arise to oppose them. First, as the interface rises, the gas volume increases, decreasing the pressure which drives the gas flow. Energy losses which are somewhat localized in the stem region occur through the generation of a secondary shock wave, fluid friction, and heat transfer to the stem walls. This also effects a pressure decrease. The principle effect is probably the ablation of mass from the stem walls which increases the mass of the flowing mixture with a resulting decrease in velocity. If the borehole diameter is relatively small, the stresses produced by the shock wave in the material surrounding the hole may be

sufficient to cause closure following the brief venting of gas resulting from the initially high interface velocities.

High-explosive simulation of an unstemmed nuclear cratering shot is complicated by the large difference in the initial interface velocities induced by the differences in initial pressures. Two approaches appear to be possible. The first involves simulating the hole of a nuclear shot with a larger HE hole. The choice of hole diameters would be based upon requiring either similar mass flows up the stem or similar cavity pressure histories. The second approach involves selecting an initial HE configuration which is similar to the shape of a nuclear cavity (including the stem) at some late time when the pressure and interface velocity have dropped to levels attainable with high explosives. Both approaches require extensive numerical modeling and detailed knowledge of all material properites affecting the gas flow. Although it will not be known with any degree of certainty until these calculations have been carried out, it appears that the above simulation approaches may not be in conflict with the previously stated criteria for fully-stemmed detonations.

(12) Summary

The effects discussed in the preceding paragraphs and the resulting similitude criteria are listed in Table 1. In the fully-stemmed configuration, most of the dynamic cratering effects can be simulated by matching the mound kinetic energies (residual particle velocities) and/or the low-pressure cavity gas adiabats. These criteria can be met individually through the judicious selection of the high explosive, the yield, and the charge configuration. Whether the criteria can be collectively met is at present unknown.

-145-

TABLE 1

SUMMARY OF SIMILITUDE REQUIREMENTS:

EFFECT	SIMILITUDE CRITERION
1. Early-time energy source behavior	None
2. Spallation mechanism	Residual mound velocities
3. Ground-shock-induced airblast	Peak surface spall velocity
4. Seismic effects	Residual mound velocities
5. Gas-acceleration mechanism	Low-pressure adiabat
6. Depth of burst and crater dimensions	Residual mound velocities and Low-pressure adiabat
7. Missile ranges	Peak surface velocity
8. Gas-vent-induced airblast	Low-pressure adiabat
9. Venting of cavity gas and radioactivity	Low-pressure adiabat
Violent venting	Particulate tracer
Nonviolent venting	? ?
10. Cloud formation	
Base surge	Residual mound velocities
Main cloud	Low-pressure adiabat (?)
11. Effects of stemming	
Stemmed case	Above criteria
Water-stemmed case	Cavity-gas/steam pressure profile (?)
Unstemmed case	Gas pressure and/or velocity profile (?)

INDIVIDUAL SIMILITUDE CRITERIA CAN BE MET THROUGH THE JUDICIOUS SELECTION OF THE HIGH EXPLOSIVE, THE YIELD, AND THE CHARGE CONFIGURATION

III. PROJECT DIAMOND ORE

(1) General Description

Project DIAMOND ORE consists of a series of high-explosive field
experiments and associated theoretical and experimental studies directed
toward (1) investigating the role of stemming materials in both nuclear
and high-explosive cratering shots at shallow and optimum depths, and
(2) arriving at specific design criteria for high-explosive simulation of
nuclear detonations. Although existing numerical codes have the capa-
bility of treating nuclear and high-explosive configurations such as those
of interest here, the codes employ many assumptions about the behavior of
matter in and about the explosive environment. The codes also require,
as input data, very detailed descriptions termed "equations of state" (EOS)
of the energy source and the cratering medium. Consequently, great
credibility cannot be attached to design specifications based on theoreti-
cal calculations alone. On the other hand, although considerable experi-
ence has been gained in the field of HE cratering, there is no means by
which this can be directly related to nuclear cratering. The intent of
DIAMOND ORE is to mesh cratering theory and experiment together by
providing a series of well-documented experiments by which the codes can
be tested thereby enhancing the credibility of the computer-generated
criteria for high-explosive/nuclear equivalence.

As presently envisioned, DIAMOND ORE is to be conducted in three
phases. Figures 3 and 4 are flow charts which depict the interactions
between the various parts of the shallow and optimum depth-of-burst programs.
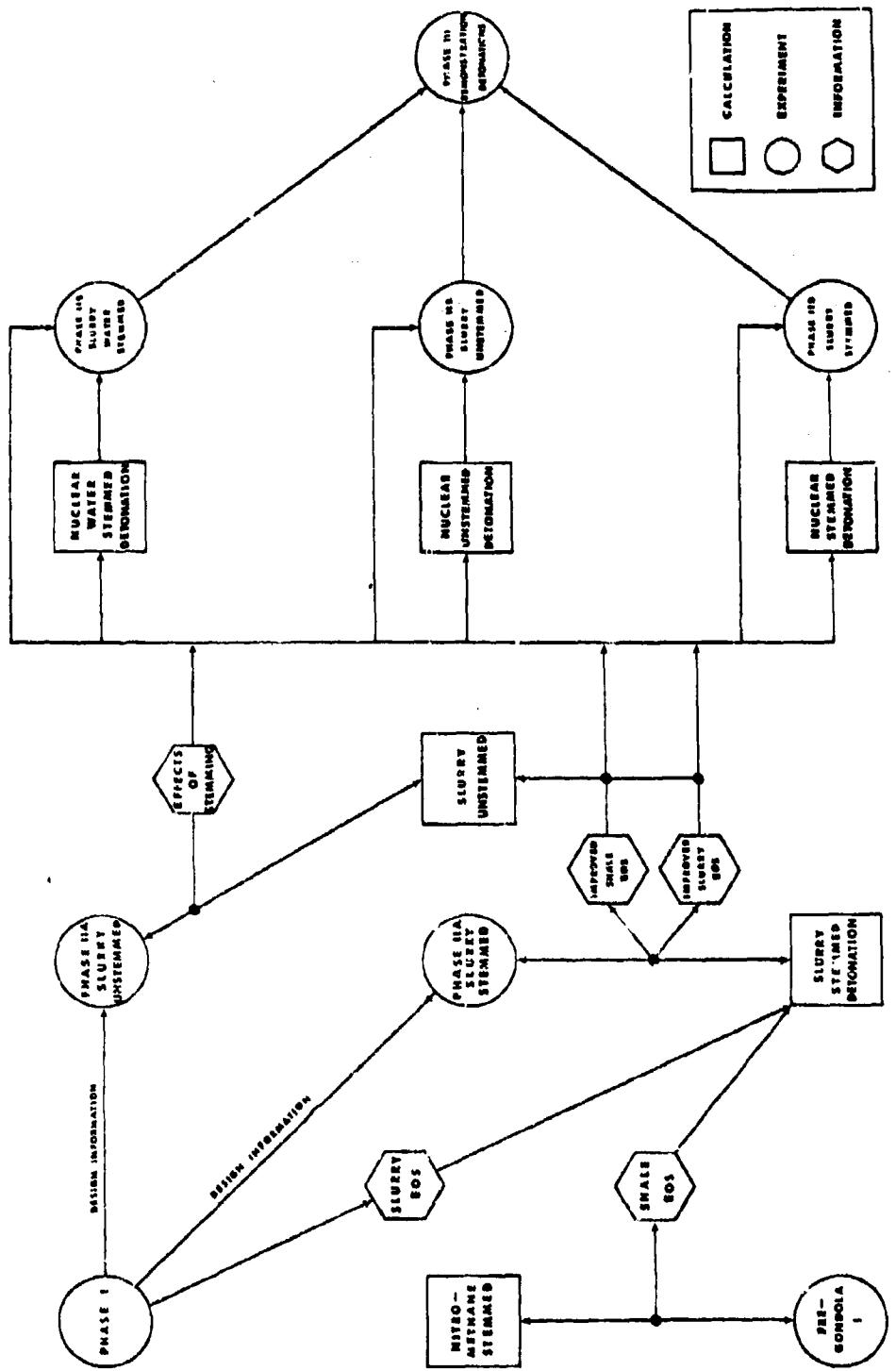
Figure 3
Diamond Ore Program at Optimum
Cratering Depth.
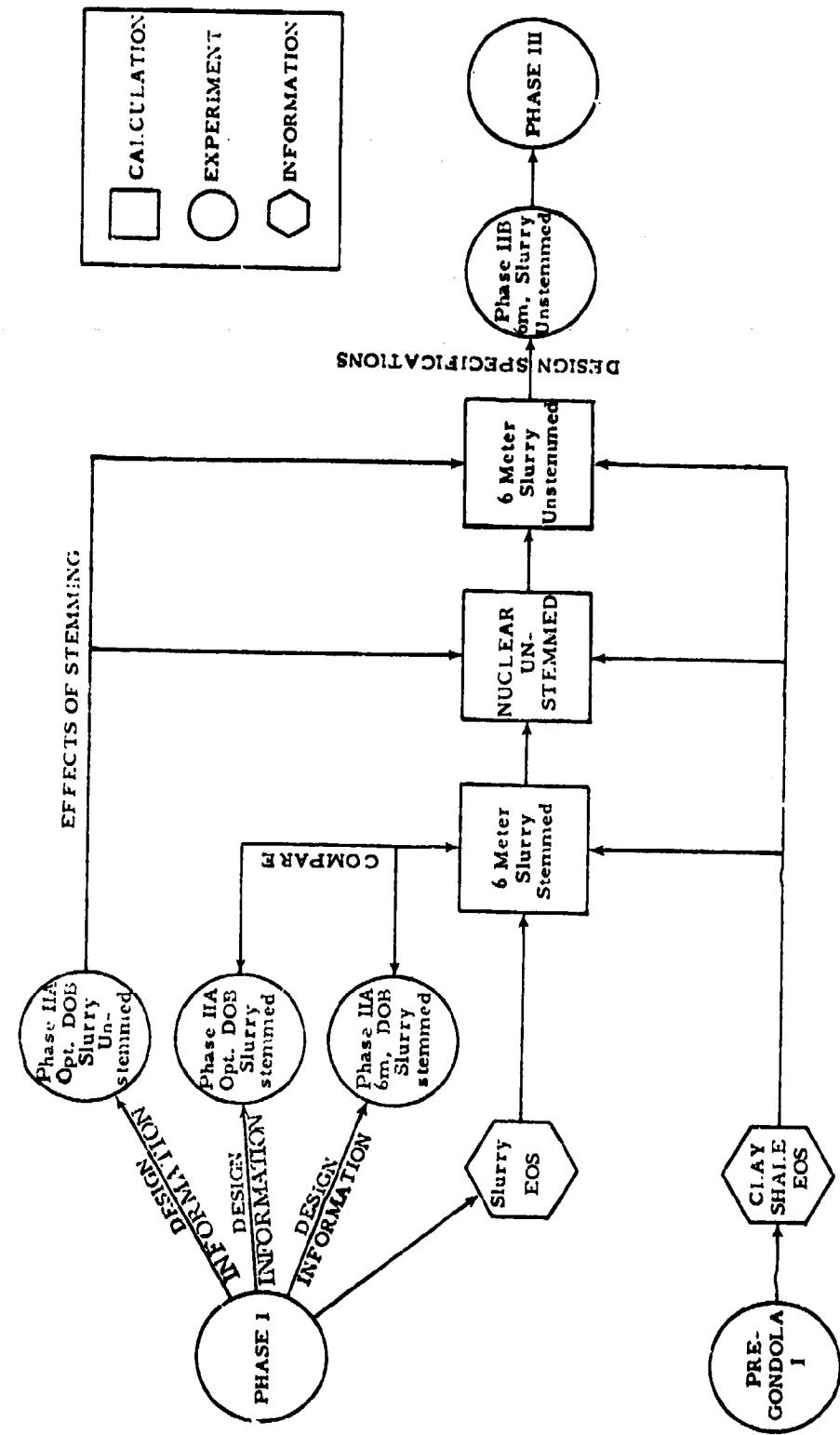
Figure 4
Diamond Ore Program at Shallow
(6 meter) Depth.

Phase I is a continuing study of high explosive characteristics, cratering media properties, and fallout simulants and has been in progress since late 1970. The fallout simulant program arising from the Phase I studies will be discussed in detail in a later section of this paper. The site of the Phase II experiments coincides with that of the Pre-GONDOLA series of 1966-68. Numerical simulation of the Pre-GONDOLA I events has been coupled with the known experimental properties of the clay shale at the site to arrive at a consistent material equation of state[8]. Similar equation-of-state studies are now being carried out on the aluminized ammonium nitrate slurry which is the explosive selected for the Phase II-A series.

Phase II includes two series of cratering detonations to be fielded at the Lewis Reservoir near Fort Peck, Montana. The first series, Phase II-A, is to be fielded in October 1971 and consists of a stemmed detonation and an unstemmed detonation at near-optimum depth of burst, and a stemmed detonation at about half-optimum depth. The yields are identical and are in the low sub-kiloton range. The results of these experiments will provide direct comparisons of the effects of stemming and depth of burst. The instrumentation will be such as to provide data on the behavior of the shale and the slurry under conditions not previously investigated. Phase II-A is then intended to serve as a well-instrumented test series upon which the hydrodynamic computer codes can be tested, and the explosive and shale equations of state refined. When the theoretical program has progressed to the point that the Phase II-A series can be correctly simulated numerically, then a thorough knowledge of the shale and slurry equations of state and the dynamic effects of stemming has been gained.

With the information gained from Phase II-A and past experience in
applying hydrodynamic codes to nuclear detonations, it will be possible to
predict the effects of nuclear detonations in the Fort Peck shale in the
following configurations: unstemmed, water-stemmed, and fully-stemmed
at optimum depth of burst, and unstemmed at half-optimum depth. Taking
into account the qualitative simulation criteria discussed in Section II of
this paper, various high-explosive emplacement configurations can be
studied to determine which will best model the nuclear explosive and its
stemming. This will be an iterative process which will converge upon the
explosive designs to be used in the Phase II-B experiments. If, at this
stage of the program, it should be determined that the slurry explosive
is incompatible with similitude requirements, the use of a more appro-
priate explosive will be considered at this time. The Phase II-B series
is scheduled for summer of 1972 and is intended to be an experimental
verification of the theoretical explosive designs.

It is anticipated that the Phase II series will provide information on
the relative importance of the various material properties as related to
stemming, depth, and yield. The design techniques employed in Phase II-B
can then be adapted to designing additional simulation experiments in other
geologic materials. This test series called Phase III would be directed
toward demonstrating the feasibility of HE simulation of nuclear explosions.

The DIAMOND ORE objectives are summarized in Table 2.

(2) Site Selection

The selection of Fort Peck, Montana, as the site of the Phase II ex-
periments was based upon several considerations. First, the clay shale
in which the series will be conducted is of interest in civil and military

-151-

TABLE 2

PROJECT DIAMOND ORE OBJECTIVES:

Design and execute high-explosive experiments which will best simulate nuclear cratering detonations under various stemming, depth, and material conditions.

PHASE 1
Study cratering media equations of state
Study high-explosive equations of state
Select high explosive for simulation
Devise fallout simulant program

PHASE IIA
Provide direct comparisons of effects of stemming and depth of burst
Provide data on behavior of clay shale and slurry under conditions not previously investigated
Serve as a well-documented experiment upon which the hydrodynamic codes can be tested and refined

PHASE IIB
Provide additional data as required
Serve as experimental verification of theoretical explosive designs in clay shale

PHASE III
Demonstrate feasibility of high-explosive simulation of nuclear cratering detonations in varied geologic media

applications. Secondly, the site is free of abundant wildlife and distance from large population centers, thereby reducing the environmental impact of the detonations. The site is sufficiently large and free of mountainous terrain to allow the implementation of a fallout simulation and collection program. The area has been the site of several other high-explosive experiments, and the experience in conducting field operations there tended to reduce project costs. However, the most important consideration was the existence of extensive data on the material properties of clay shale, thereby reducing data acquisition costs.

(3) Explosive Selection

In fielding an experimental cratering program as extensive as DIAMOND ORE, the cost of explosives becomes a primary consideration. Past experience in the field of explosive excavation has shown that, in terms of overall cratering efficiency, the slurry explosives are by far the most economical. Because slurry can be emplaced in liquid form and is classed as a blasting agent rather than a high explosive, costs of handling are also reduced.

As will be discussed later, the use of a fallout simulant requires that the simulant particles be uniformly distributed throughout the HE. Because the slurry can be made to gel after emplacement, it is ideally suited to this purpose. The simulant particles can be uniformly mixed while the slurry is liquid and will remain so after emplacement.

Finally, some slurry formulations provide very high energy release per initial unit volume, thereby more closely approximating the initial nuclear cavity than less energetic explosives. Unfortunately, the slurry equation of state is presently unknown, and the more precise similitude

-153-

requirements discussed in Section II of this paper therefore cannot be applied to it. In view of its other desirable properties and the fact that it is of considerable interest in its own right, an attempt to make use of slurry as a nuclear explosive simulant was felt to be justified. Accordingly, slurry was selected for the Phase II-A experiments; and determining its equation of state was made one of the principal objectives of Phase I. However, should the slurry prove unacceptable for simulation purposes, concurrent studies are being conducted to select an alternative.

(4) Phase II-A Technical Programs

The following outlines the technical programs to be executed in conjunction with Phase II-A.

(a) Site Investigation. Prior to the final selection of the ground zeros for the three Phase II-A events, seismic refraction surveys and continuous borehole loggings were conducted at a number of proposed sites. These investigations provide direct measurements of sonic velocities and indications of anisotropies in the shale.

(b) Direct Measurements of Material Properties. Core samples have been provided to the U. S. Army Engineer District, Missouri River Division, for measurements of unconfined compressive strengths, residual shear strength, Young's modulus, Attenberg limits, and other material properties. Samples have also been provided to L^3 for plane strain measurements which will yield data on the static elastic moduli and material strength.

(c) Close-in Ground Motion and Earth Stress Measurements. (Figure 5) For the two optimum DOB detonations, acceleration, velocity, and stress gauges will be employed in a linear array of boreholes to provide the
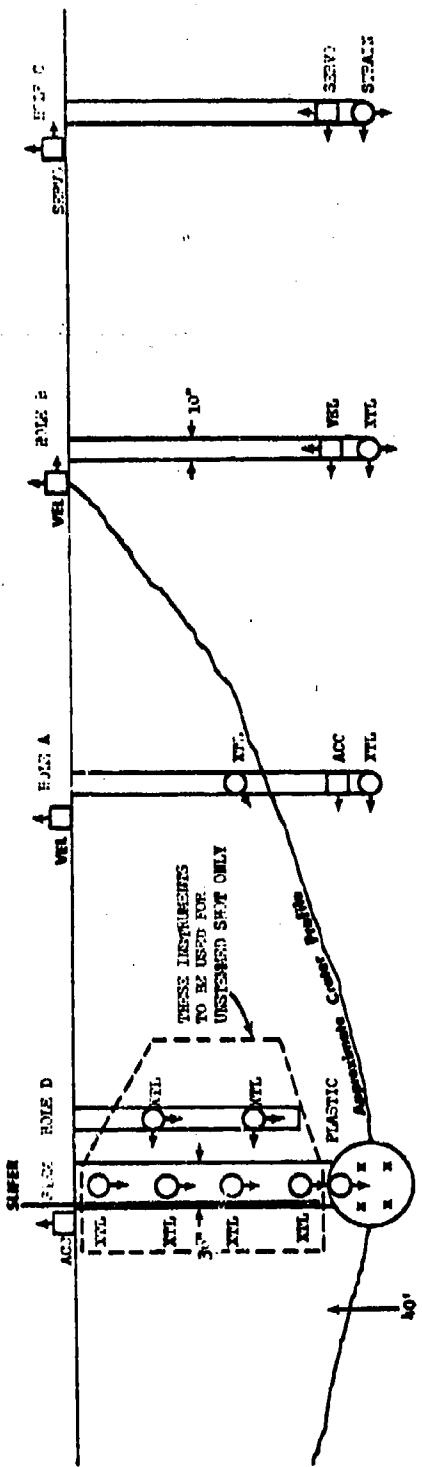
-154-

Figure 5

Instrumentation layout for Phase IIA of Project Diamond Ore. The gauges designated are used only on the unstemmed event. Only the surface gauges, slifer, plastic gauge, and rate sticks are used on the 6-meter event.

motion and stress time histories of the medium under the influence of the shock front for use in L^3 and S^3 postshot calculations. It is anticipated that the radial and tangential gauges in holes B, C, and D may provide data for determining dynamic values for the elastic moduli.

(d) Surface Motion Measurements (Figure 5). Early time high-speed photography (1000 FPS) will record each detonation for the study of velocities and the acceleration of particles within the crater area, and for the study of mound growth and venting. Accelerometers and velocity gauges will also be used to measure the rate of mound growth.

(e) Stemming Pressure Measurements (Figure 5). A vertical array of stress gauges will be emplaced within the air column of the unstemmed shot. An additional array will be located just outside the unstemmed hole. Both arrays will provide stress and time-of-arrival measurements required for code calculations.

(f) Uniformity of Detonation Measurements (Figure 5). The detonation history of the explosive detonation will be recorded by emplacing four "rate sticks" within the explosive material. These measurements are required to verify the detonation uniformity of the slurry explosive and will provide detonation velocity and perhaps shock velocity data for the slurry.

(g) Crater Measurements. The dimensions of all craters will be measured by conventional ground survey to provide preliminary crater dimensions. Aerial mapping will provide precise dimensions of experimental results. Aerial photography of the craters will be taken after all three detonations have been conducted.

(h) Ejecta Analysis. Missile ranges and directions relative to each ground zero will be determined by conventional ground survey for each

detonation. This data will be used to relate missile areal densities to missile ranges in the area between the limits of the continuous ejecta and the maximum missile range. Areal distribution of the continuous-ejecta will be determined by the "point count technique". Limits of the continuous ejecta and maximum missile range will be marked in a manner which will be detectable on the postshot aerial photography.

(i) Technical Photography. In addition to the photography mentioned above, documentary 16mm photography will record the cloud formation for each event, and incidental still photographs will be taken of each detonation. Although no cloud studies are planned at this time, the information will be available should such studies be desired at a later date.

(j) Fallout Collection. Fifty (50) collection trays will be placed in the expected downwind area on all three detonations in order to provide basic data as to fallout mass deposition relative to ranges from Ground Zero. This data is desired for the fallout collection array designs to be used in Phase II-B. Samples from the fallout trays will be taken for use in tracer background studies. The location of these trays relative to the wind direction at shot time will not be a controlling factor in the firing decision. Although no fallout simulation is scheduled for Phase II-A, uncoated silica particles will be mixed with the explosive.

(k) Meteorological Program. Wind speed and direction will be determined three times a day (at 0900, 1200, and 1500 hrs) at the ground surface and at altitudes of 100, 200, 300, 400, 500, 600, 800, 1000, 1500, 2000, and 2500 meters. This data collection will be initiated five days prior to shot date for Event II-A-1 (unstemmed) on a daily basis. In addition, on shot days, wind data collection will be initiated at the time

of detonation.

(1) Close-in Airblast. Peak overpressures within the range of 0.03 to 1.0 psi will be measured on all three shots at ranges of 250', 500', 1000', and 2000'. Additional gauges will be at 4000' on the unstemmed shot shot and at 4000' and 10000' on the half-optimum depth shot.

(m) Seismic Investigation. Two ground-shock stations will be employed to measure particle velocity for comparative seismic measurements using EERO in-house capability. Gauge locations will be field determined and surveyed after emplacement.

IV. VENTED RADIOACTIVITY AND FALLOUT DISTRIBUTION

To determine the effect the type and degree of stemming will have on the fraction of material vented from the simulated nuclear crater and its distribution around the shot point, tagged mineral particles are being employed which can be identified in the post-explosion debris. The particles are of a size and density such that, while in the debris cloud, they will respond to the influences of winds and gravity as do the same size radioactive particles from a nuclear explosion. Thus, if the venting process of a nuclear explosive is simulated, particle areal deposition should be duplicated for a similar wind condition. The nuclide composition of a particle and the resulting fallout field dose rate will still require generation by numerical techniques.

(1) Previous Fallout Simulation Experiments

While the use of tracers for this purpose is not new, the approach being taken is new. In previous experiments, Project YO-YO and Pre-BUGGY, radioactive tracers (^{198}Au and ^{76}As in Project YO-YO and ^{140}La in Pre-BUGGY) were employed. Conclusions from these experiments were that: (1) The fraction of the radioactive material that escaped was related to depth of burst and apparent crater radius, but the escape fractions were higher than those expected from nuclear detonations at the same ratio of depth of burst and apparent crater radius; and (2) radioactive tracers in chemical explosives did not simulate fallout.

Examination of these experiments indicates that the method of placement of the radioactive tracer in the chemical explosive may have lead to the relatively high amount of radioactivity vented. In both experiments, the radioactivity was placed in a capsule located in the center of the charge.

Upon detonation there was no assurance that, prior to venting, the radio-active tracer would uniformly mix with the explosive gases or would not be preferentially vented in a gaseous jet.

Failure of the radioactive tracer to simulate fallout is attributed to the differences in the temperature, pressure and venting processes between nuclear and chemical explosives. As was previously noted, the venting process is strongly influenced by the cavity pressure and temperature history and, in turn, the fraction of volatile, refractory and intermediate radionuclides vented. Summarized, it is known that if the temperature and pressure at vent time are higher, a more violent mixing of radionuclides will result and the vented material will contain a higher fraction of refractory and intermediate radionuclides. Distribution of the refractory and volatile radionuclides differs among fallout particle sizes with the cavity pressure and temperature at vent time. When pressures and temperatures are _low_, refractory nuclides are deposited on the larger particles of overburden and fall close to ground zero. When pressures and temperatures are _high_, refractory elements are associated with particles having a smaller mean radius and are more abundant in the nuclear cloud. Then, refractory radionuclides are found to be volume contained while volatile elements are surface deposited.

Data from the Pre-BUGGY experiment[9] showed that the radioactive tracer employed (^{140}La) was surface adsorbed on the particles in the detonation cloud. For Pre-BUGGY, 96% of the recovered radioactivity was found on small particles which constituted only 8% of the particle mass but an estimated 90% of the surface area of the deposited particles. With the majority of tracer activity associated with the smaller particle

sizes, winds acting on the explosion-produced cloud carried the small particles to great distances from surface zero and distorted the fallout patterns.

To circumvent the problems of previous experiments, the approach being pursued to measure the vent fraction consists of using a neutron activable chemical element surface adsorbed on quartz particles of known size and density. These tagged particles are uniformly mixed throughout the explosive charge. The rationale for this approach is discussed in the following paragraphs.

(2) Tracer Selection

Since the test program is to be conducted in a number of locations and geologic medium, an inert tracer, suitable for neutron activation after collection, is being employed to avoid the problems associated with the use of radioactive materials. Use of neutron activable tracers rather than a radioisotope eliminates:

(a) The need for an AEC or State license for use of radioactive materials.

(b) The logistic problems of providing sufficient tracer for a field test that is weather dependent.

(c) The requirement to have field equipment to measure radioactivity levels.

(d) The difficulties of converting radiation measurements to meaningful fallout values.

(e) The radiation control measures required during handling of the tracer and placement in the explosive charge.

(f) The problems associated with releasing radioactivity to the

-161-

environment.

The neutron activable tracer coated on particles of known size and density is being mixed with the explosive to assure a uniform mixture in the gases resulting from the explosion and to prevent preferential venting of the tracer.
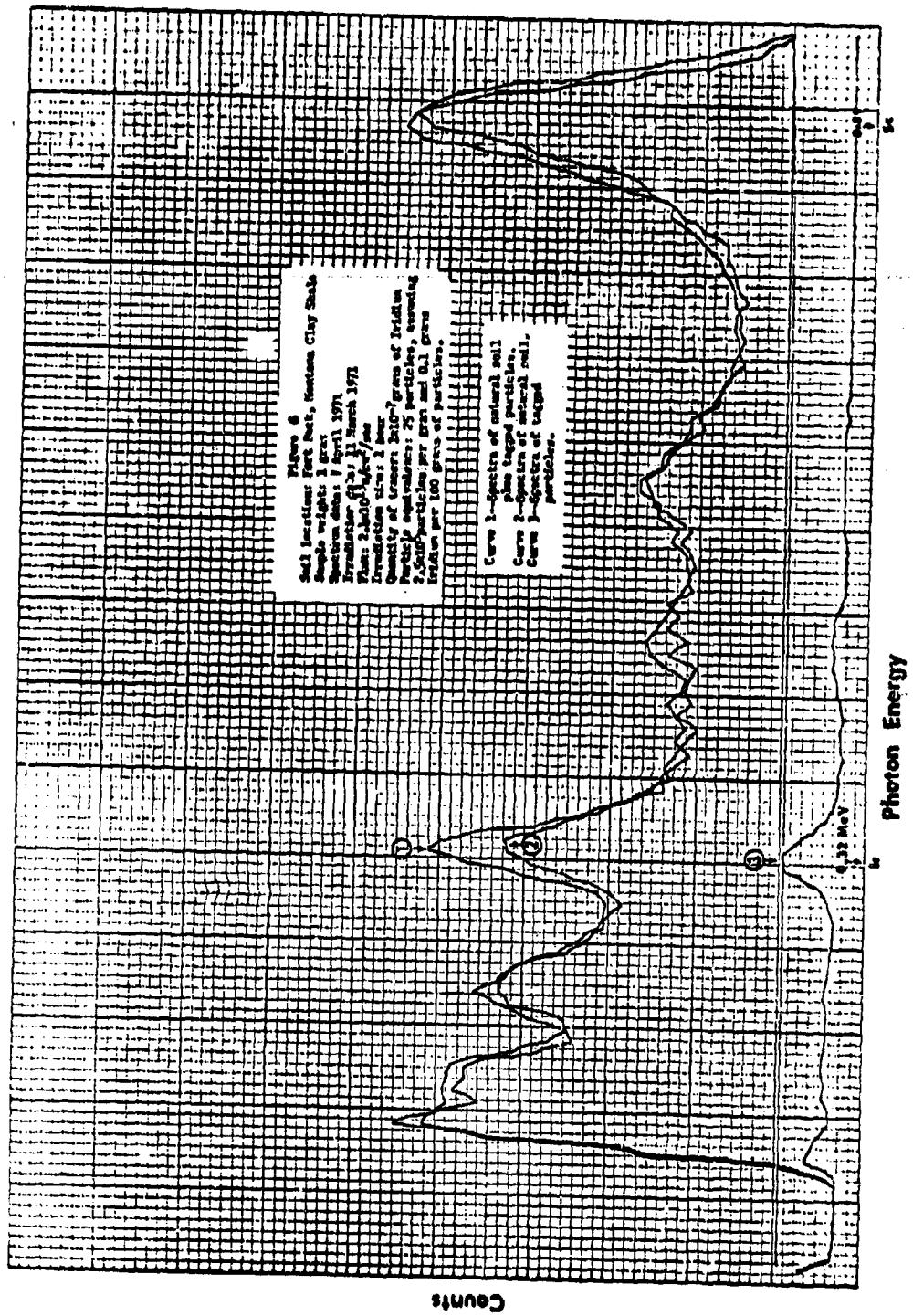
The coated particles range in size from 125 to 175 microns and are a pure quartz (Wedron sand) with a particle density of 2.65 grams/cc. This size range was selected because it comprises an important size fraction in local fallout and falls to earth within a reasonable distance from ground zero. This limited downwind fallout area permits an extensive fallout sample collection effort to be fielded. This particle range also brackets the 143-micron particle size assumed to be nominal in certain fallout prediction schemes. Quartz particles are being employed since quartz has a high melting point (1600°C) and will survive the high explosive environment without melting. In addition, when irradiated, pure quartz produces no radionuclides which interfere with the detection of the tracer selected.

Iridium is the neutron activable element of choice for coating the particles. This element was selected after reviewing the literature[10, 11] pertaining to the elemental constituents of various geologic media, investigating the physical properties of various chemical elements suitable for neutron activation, and examination of the gamma ray spectra of geologic media from proposed test site locations. The literature review indicated that the natural abundance of iridium in geologic media was about 1×10^{-9}. This low abundance in nature minimizes the interference with detection of traced particles from naturally occurring iridium. Iridium's melting

point is approximately 2400°C. This high melting point minimizes the loss of iridium from the coated particles due to boiling. The explosive temperatures are estimated to be about 2500°C. The thermal neutron cross section of ^{191}Ir which has a natural abundance of 37.3% is 750 barns. Its activation product (^{192}Ir) has a half-life of 74 days and undergoes beta decay to ^{192}Pt, emitting 316 kev gamma photons with a branching intensity of 85% and 468 kev photons with 50%. The high cross section permits activation in relatively short periods in the reactor (1 hour at a neutron flux of 5×10^{12} n/cm^2/s) and the high abundance of photons permits easy detection. Spectra of various irradiated media were examined in a 3" x 3" sodium iodide crystal and no naturally occurring iridium of significance was found to be present. After 20 days, the gamma photons present from other long-lived radionuclides present only minimal interference with the detection and quantification of the iridium. Figure 6 shows the spectra of an irradiated medium from a site. In the figure line 2 is the natural spectra of the soil; line 3 shows the iridium spectra; and, line 1 shows the spectra of 10^{-7} grams of iridium (approximately 25 coated particles) in one gram (approximately 2×10^5 particles) of natural soil after sieving to obtain the 125 to 175 micron particles.

(3) Preliminary Simulation Tests

In addition to the studies indicated above, a series of tests have been conducted to determine if the iridium-coated quartz particles will survive an explosive environment. In one set of tests, 0.4 pounds of iridium-coated quartz particles were uniformly mixed with 4 pounds of C-4 explosive. These explosive charges were detonated 2.4-feet below the ground surface and "fallout" samples collected and sieved into three fractions:

Figure 6

Soil location: Fort Peck, Bearpaw Clay Shale
Sample weight: 1 gram
Spectrum date: 1 April 1971
Irradiation date: 11 March 1972
Flux: 2.1x10^13 n/sec/cm²
Irradiation time: 1 hour
Quantity of iridium: 3x10^-7 grams of iridium
Density of sedimentation: 25 particles, covering
2.5x10^-8 particles per gram and 0.1 grams
iridium per 100 grams of particles.

Curve 1—Spectra of natural soil
          plus target iridium,
Curve 2—Spectra of natural soil.
Curve 3—Spectra of target soil
          particles.

Photon Energy

Counts

-164-

(1) less than 125 microns, (2) 125 to 175 microns, and (3) 175 to 350 mi microns. Analysis of the collected samples indicated that the iridium did not appreciably migrate from the 125 to 175 micron range.

In another set of 3 tests, 0.8 pounds iridium-coated particles were uniformly mixed with 8 pounds of alluminized ammonium nitrate slurry called TD-2 and manufactured by the IRECO Company of Salt Lake City, Utah. The results of the three tests were identical and indicated that the fallout could be collected and the iridium content or particle content could be determined after irradiation.

Cratering tests performed with TD-2 explosives containing various weights of sand per weight of explosives indicated that explosive performance, for cratering purposes, was not seriously degraded if the amount of sand added to a charge was not more than 10% of the charge weight.

To assist in the transition from pound-charge tests to multi-ton charge field events, a series of four 1-ton charges containing 10% iridium-coated sand was executed at Trinidad, Colorado, during the period 9 to 25 September 1971. All four charges were emplaced in a massive sandstone medium at a constant depth of burial, 20 feet, but with varying stemming conditions: full stemming, water stemming, no stemming in a 36-inch diameter hole; and no stemming in a 4-inch diameter hole.

(4) Tracer Collection

To determine the amount of material vented from each of the Trinidad tests, approximately 200 debris samples were collected about each test. Debris was collected in 24" x 24" x 2" aluminum trays. Each tray contained louvers spaced one inch apart and slanted at 45 degrees to the horizontal to insure that particles arriving at a tray did not bounce or

blow out of the tray collector. Figures 7 and 8 show the "fallout" collector array which extended out to 2000 yards from the surface zero of each detonation.

(5) Selection of Tray Locations

Selection of collector tray locations was based on estimates of: (1) the limit of continuous ejecta, (2) a mass deposition equation developed by M. Nordyke of L^3 (unpublished), (3) fall times of tagged particles from the top of the cloud, (4) detectability of mass deposition measurements, and (5) cloud data from previous 1-ton detonations conducted at the Trinidad Test Site.

Nordyke's equation is:

$$\delta = \frac{5 \times 10^6}{1 + 0.02 \, v^3} \left(\frac{R}{W^{1/3}}\right)^{-(3.85 \pm 0.1v)}$$

where δ = mass deposition (gm/m^2),

  W = yield of the detonation (kt),

  v = wind speed (mph),

  R = distance from SGZ (10^2 yds),

The plus in the exponent is for upwind and crosswind predictions, while the minus is for downwind predictions.

There are several restrictions for the use of the equation, some of the major ones are:

a.   for downwind and crosswind predictions, $10 \leq \delta \leq 10^4$;

b.   for upwind predictions, $10^2 \leq \delta \leq 10^4$

c. $v \leq 20$ mph.

Figure 9 shows the predicted mass deposition versus downwind distance for the 1-ton Trinidad detonations and indicates the apparent crater radius
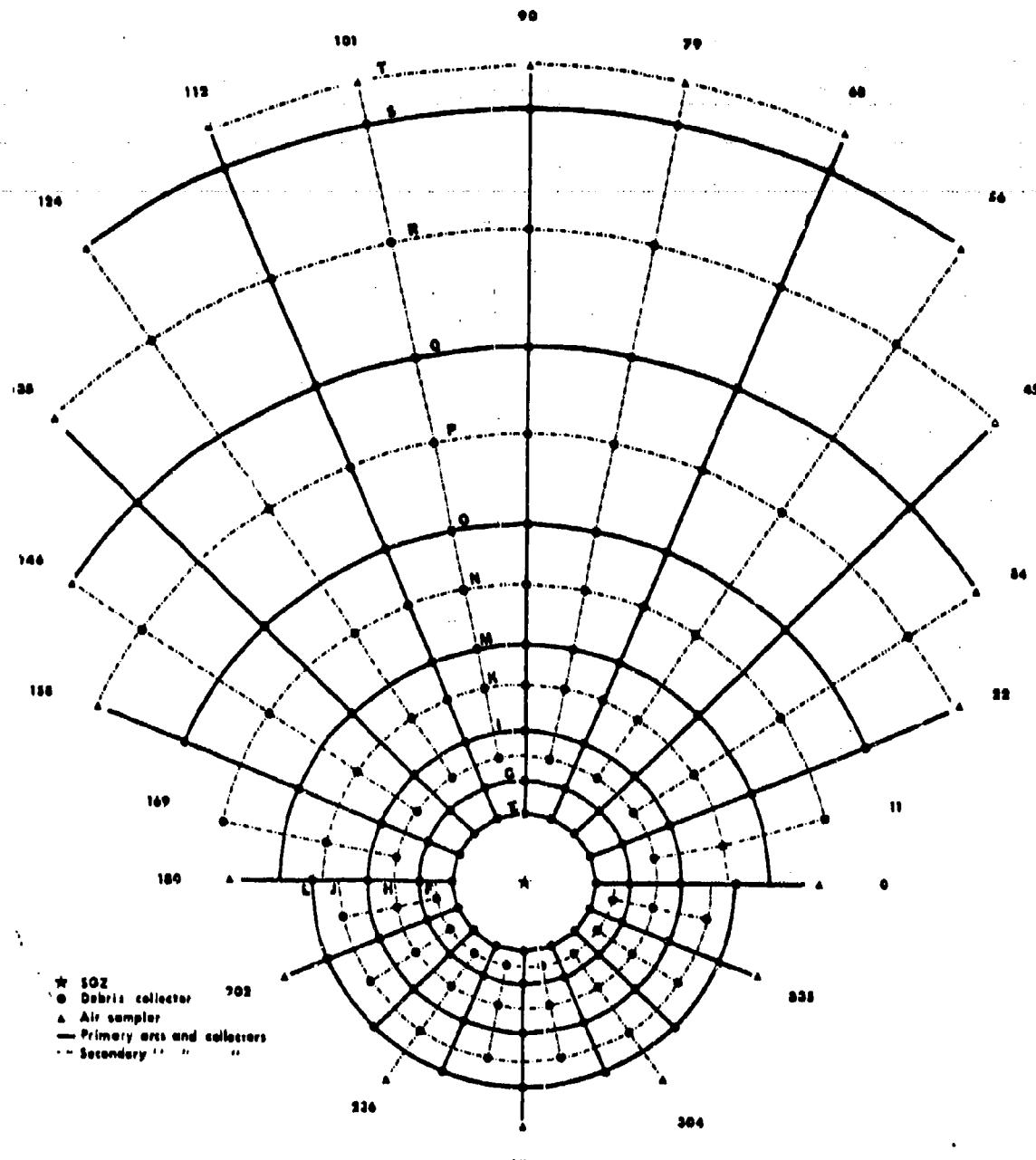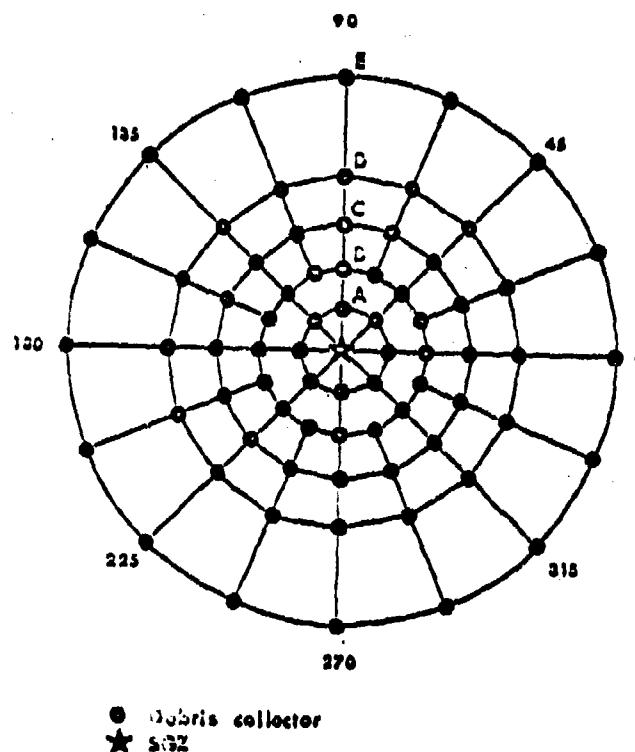
Figure 7
Intermediate Range Collector
Array for Project Trinidad.

Figure 8
Close-in collector array for
Project Trinidad with ranges.

ARC DESIGNATION	DISTANCE FROM SGZ (yds)
A	25
B	50
C	75
D	110
E	170
F	210
G	250
H	310
I	370
J	440
K	480
L	500
M	560
N	730
O	800
P	1100
Q	1300
R	1600
S	1900
T	2000
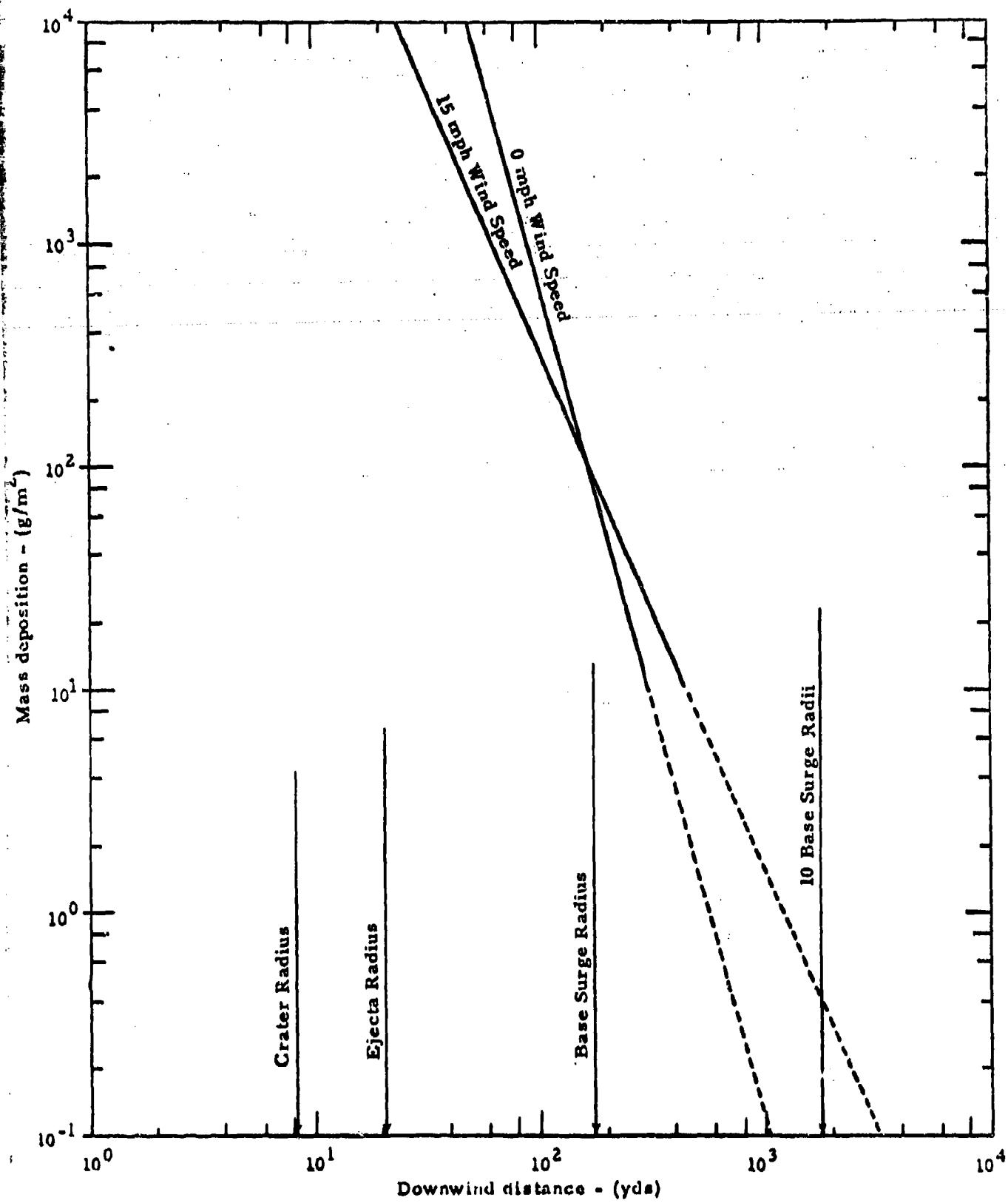
Figure 9. Predicted mass deposition versus downwind range for Trinidad.

and limits of continuous ejecta for previous 1-ton detonations in the sand-
stone medium.

The detonation times were selected so as to constrain the wind to a
speed of 5 to 15 miles per hour centered along the 90 degree axis of the
array. Cloud height, growth, and travel photography and wind velocity
measurements at each 100 ft up to 2000 ft were taken to provide data for
the subsequent analysis to be performed.

(6) Data Analysis

The collected material from each tray was weighted to determine the
mass deposition per unit area and then sieved into size fractions: 88 to
125 microns, 125 to 175 microns, and 175 to 350 microns for irradiation.
These three samples from each collection point will be encapsulated,
irradiated with thermal neutrons, and the number of tagged particles
determined by measuring the ^{192}Ir content. The size fractions over and
under the tagged particle size will also be processed to verify that agglom-
eration or fracturing of the particles did not occur, and that the iridium
did not migrate.

Data from the mass deposition, sieve fractions, and ^{192}Ir content
will permit an integration over the area in the fallout pattern to compute
a fraction vented for each of the events.

The fallout pattern from the fully-stemmed high-explosive detonation
will be converted to radiation dose-rate contours by relating to fully-
stemmed nuclear detonation data. Data from the 4-inch and 36-inch un-
stemmed and the water-stemmed detonations will be similarly treated
using best estimates for the fractionation effects which may occur. Frac-
tionation effects for these cases are the subject of additional theoretical
study.

ACKNOWLEDGEMENTS

REFERENCES

1. T. R. Butkovich, *The Gas Equation-of-State for Natural Materials*, LRL Rept., UCRL-14729 (1967).

2. J. B. Knox, et al, *Radioactivity Released from Underground Nuclear Detonations: Source, Transport, Diffusion and Deposition*, LRL Rept. UCRL 50230 Rev. 1 (1970).

3. Y. A. Izrael and E. D. Stukin, *The Gamma Emission of Radioactive Fallout*, AEC-tr-7100 (1970).

4. R. F. Bourque, *Project TRENCHER: Evaluation of Aluminized Blasting Agents for Cratering and Hole Springing*, EERO Rept. TR-28 (to be published).

5. D. J. Fitchett, *MIDDLE COURSE Cratering Series*, EERO Rept. TR-35 (to be published).

6. A. G. Gaydon and I. R. Hurle, *The Shock Tube in High-Temperature Chemical Physics*, Reinhold Publ. Corp., New York (1963).

7. B. K. Crowley, et al, *An Analysis of Marvel - A Nuclear Shock-Tube Experiment*, LRL Rept. UCRL-72489 (1970).

8. B. K. Crowley, D. E. Burton and J. B. Bryan, *Bearpaw Shale: Material Properties Derived from Experiment and One-Dimensional Studies*, LLL Rept. UCID-15915 (to be published).

9. W. B. Lane, et al, *Use of ^{140}La as a Radiotracer in Chemical Explosives*, USNRDL TR-638 (1963).

References, continued

10. W. C. Day and R. A. Paul, Trace Elements in Common Rock Types and Their Relative Importance in Neutron-Induced Radioactivity Calculations, NCG TM 67-7 (1967).

11. H. P. Yule, "Experimental Reactor Thermal-Neutron Activation Analysis Sensitivities" in Analytical Chemistry, Vol. 37, No. 1 (1965).
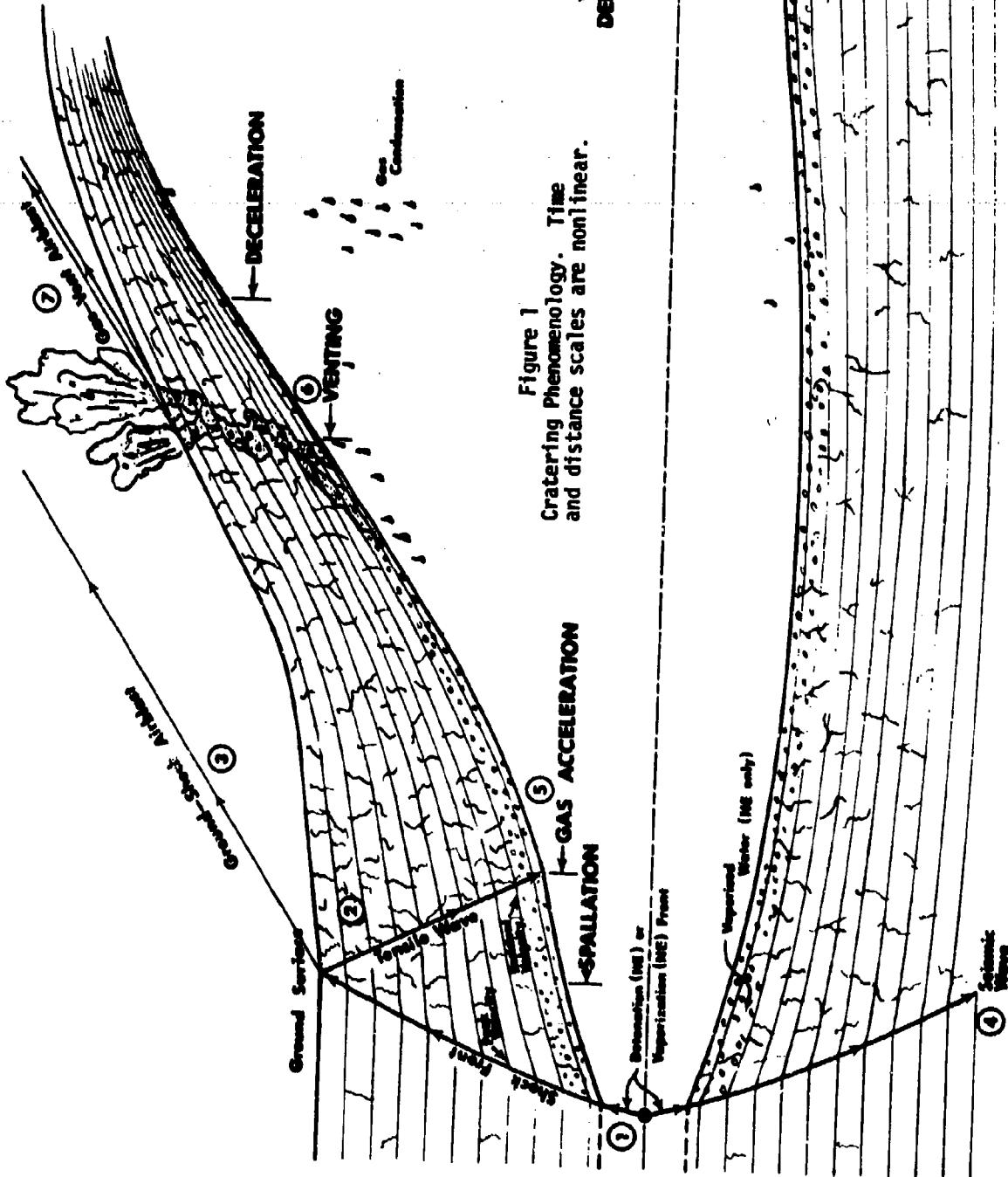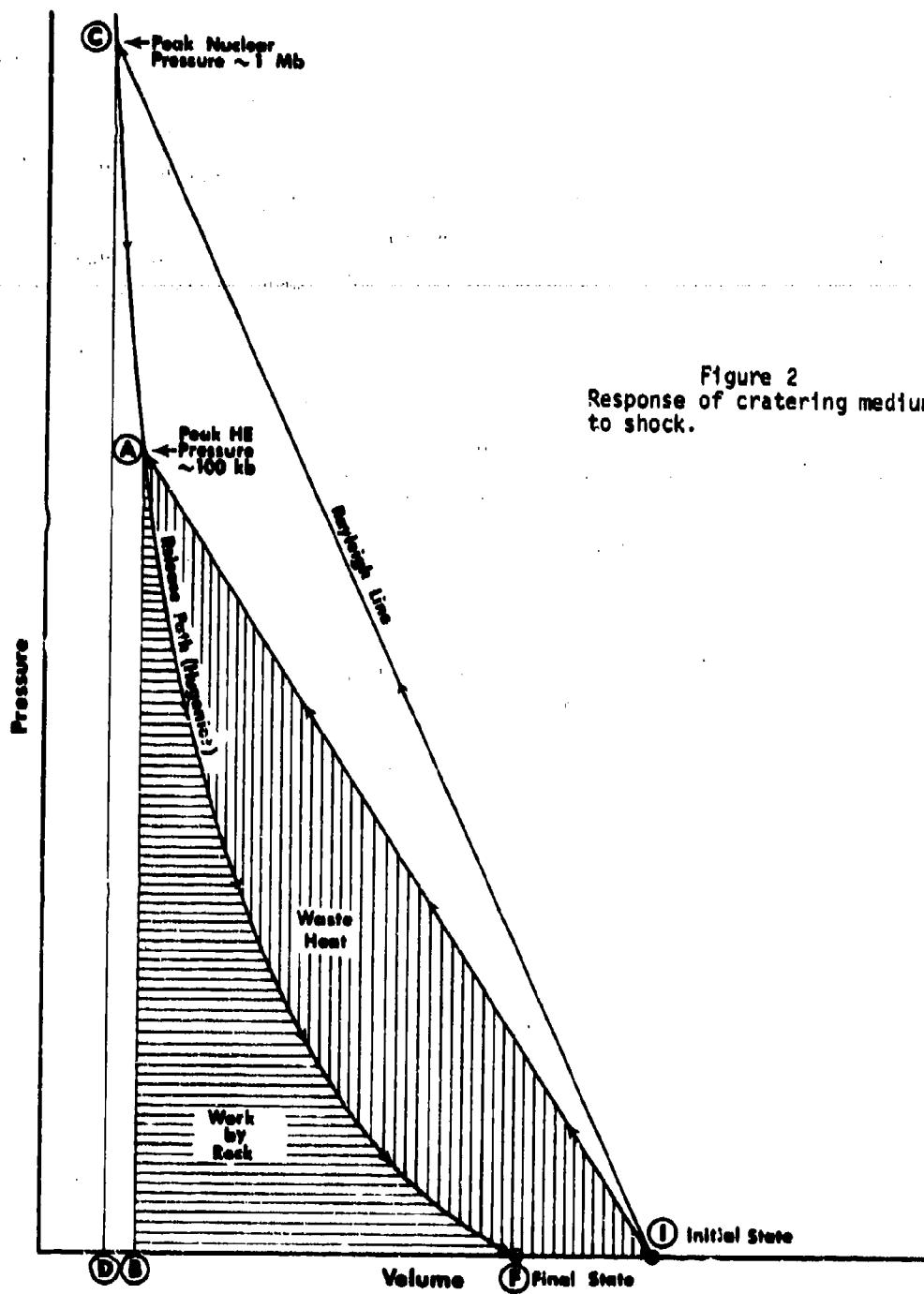
Figure 1
Cratering Phenomenology. Time and distance scales are nonlinear.
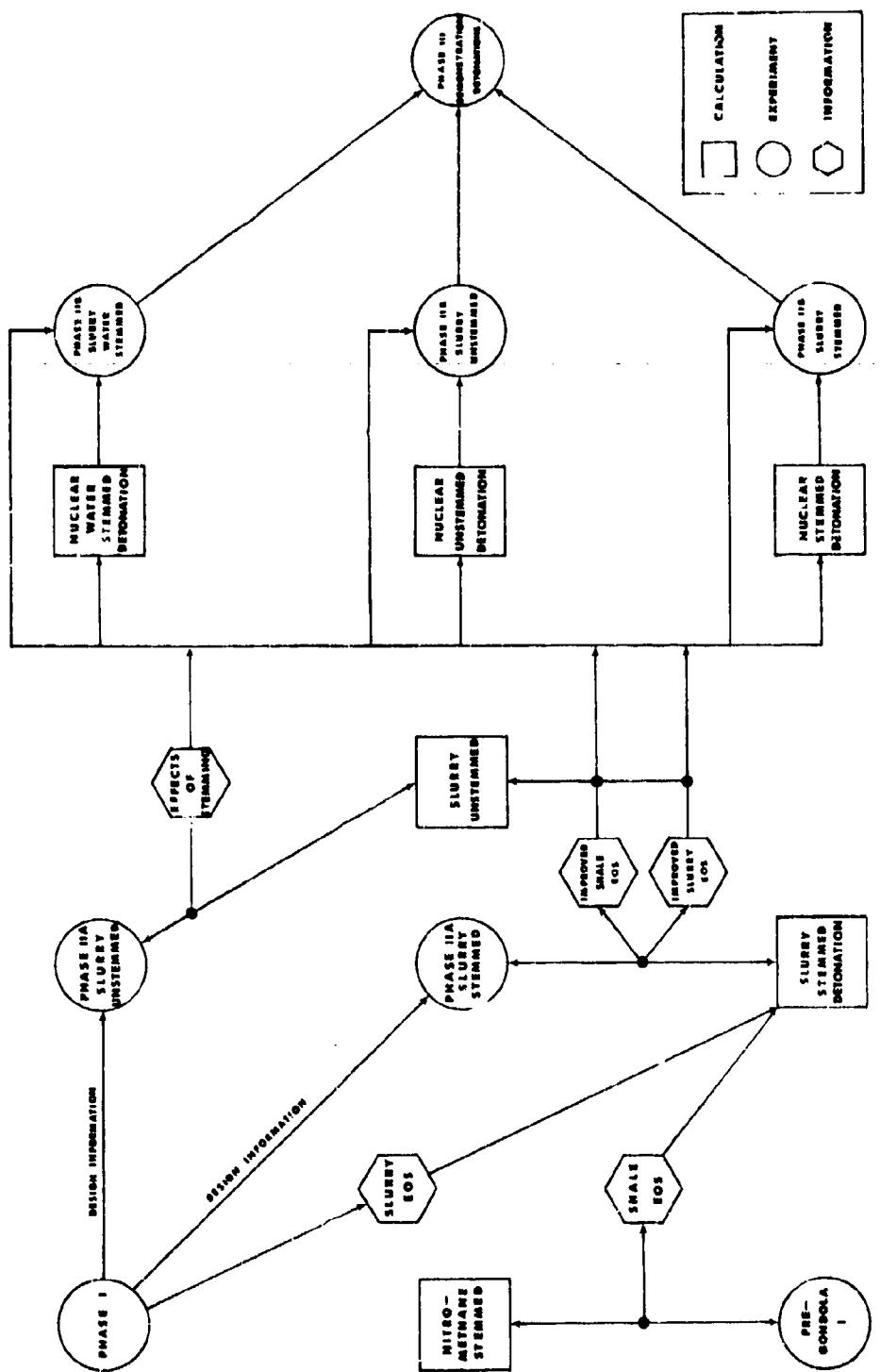
Figure 2
Response of cratering medium
to shock.

Figure 3
Diamond Ore Program at Optimum
Cratering Depth.
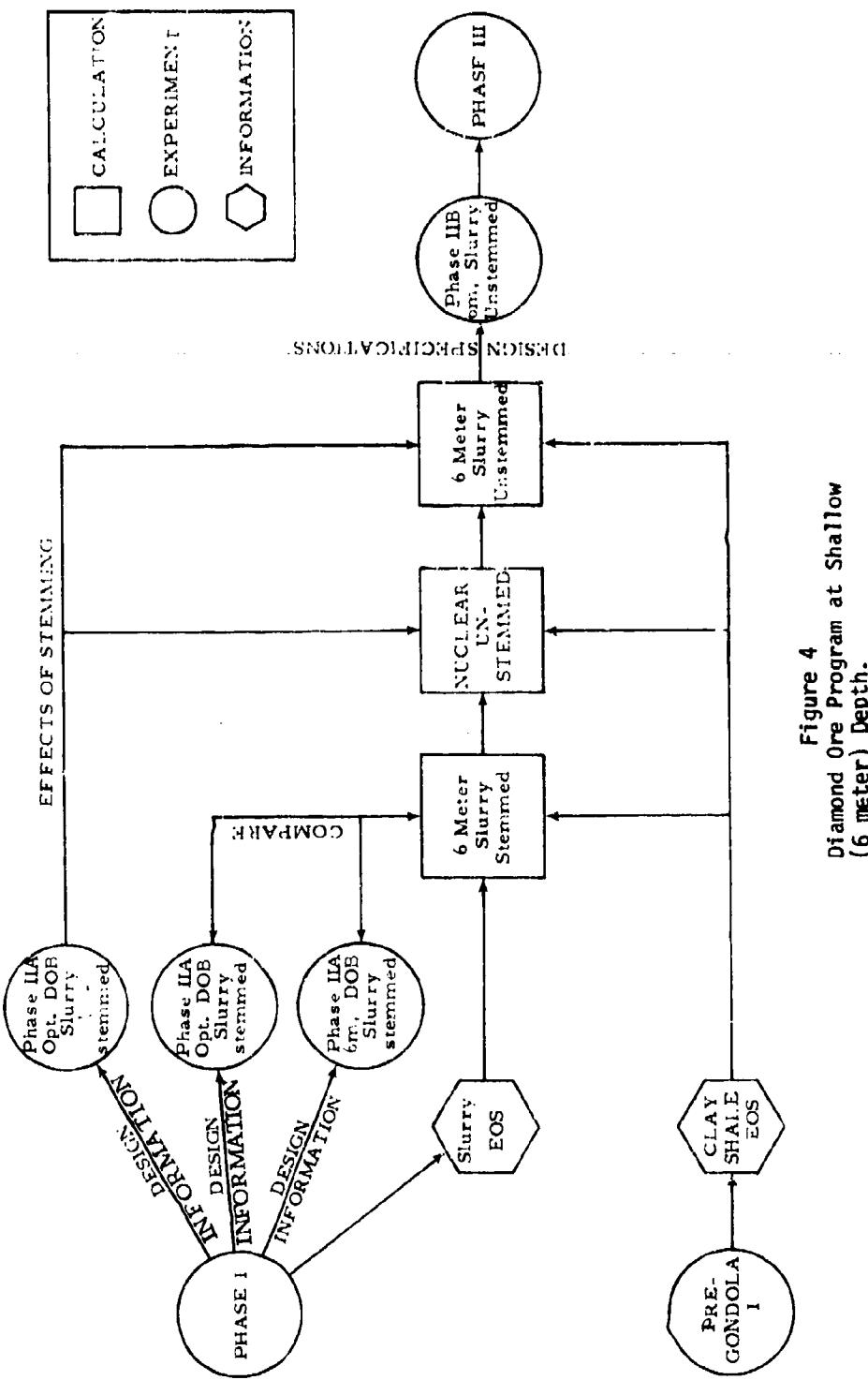
Figure 4
Diamond Ore Program at Shallow
(6 meter) Depth.

Figure 5

Instrumentation layout for Phase IIA of Project Diamond Ore. The gauges designated are used only on the unstemmed event. Only the surface gauges, slifer, plastic gauge, and rate sticks are used on the 6-meter event.

Figure 4

Soil location: Fort Peck, Montana Clay Shale

Sample weight: 1 gram
Spectrum date: 7 April 1971
Irradiation date: 11 March 1971
Flux: 2.4x10⁻³ n/cm²/sec
Irradiation time: 1 hour
Quantity of tracer: 1x10⁻⁷ grams of Iridium
Particle equivalence: 25 particles, assuming 2.5x10⁵ particles per gram and 0.1 gram Iridium per 100 grams of particles.

Curve 1—Spectra of natural soil plus tagged particles.
Curve 2—Spectra of tagged particles.
Curve 3—Spectra of natural soil.

Photon Energy

Counts

-179-

Figure 7
Intermediate Range Collector
Array for Project Trinidad.
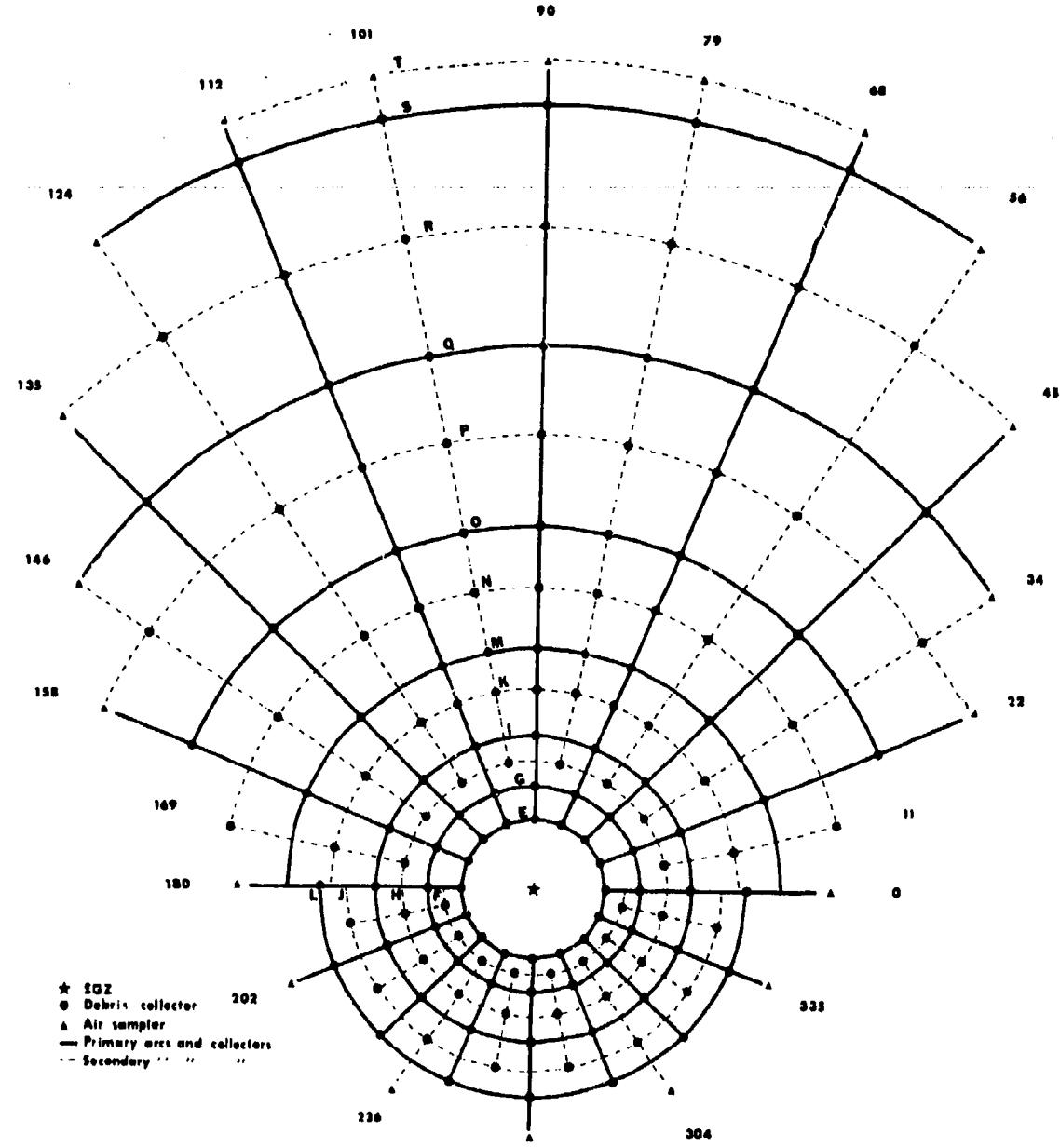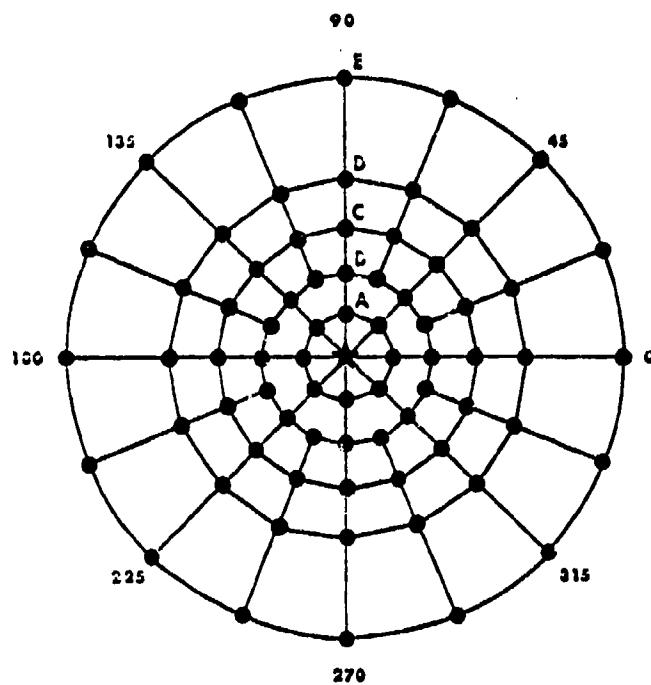
Figure 8
Close-in collector array for
Project Trinidad with ranges.

ARC DESIGNATION	DISTANCE FROM SGZ (yds)
A	25
B	50
C	75
D	110
E	170
F	210
G	250
H	310
I	370
J	440
K	480
L	500
M	580
N	730
O	880
P	1100
Q	1300
R	1600
S	1900
T	2000

● Debris collector
★ SGZ
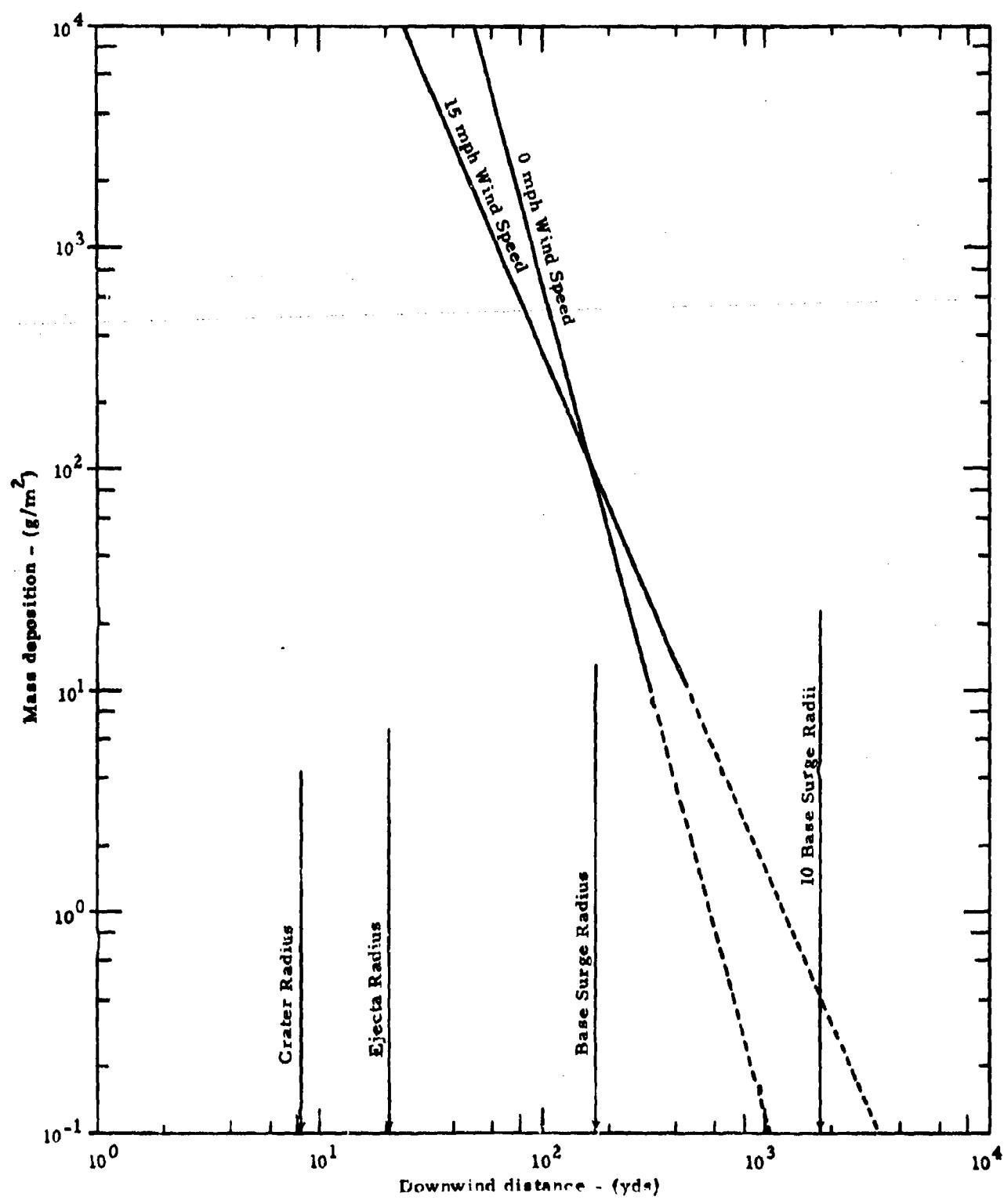
Figure 9. Predicted mass deposition versus downwind range for Trinidad.

MODIFIED FACTORIAL EXPERIMENTS
FOR ANALYSING POISSON DATA

L. Ott and W. Mendenhall
University of Florida

I. Introduction

Recent emphasis on research in the areas of air and water pollution has focused attention on the need for better methods of analysing Poisson data. We refer particularly to the problem of relating mean response to a set of independent variable for purposes of estimation. Thus an experimenter may wish to estimate the mean particle count per unit volume of air at a given location as a function of the rates of input of pollutants, wind velocity and direction. Or, he might wish to relate the accident count per unit time to a set of causative or related variables.

Statistical analyses of the effect of a set of independent variables on a Poisson response have traditionally relied on the use of a transformation on the response data. Thus a square root (or modified square root) transformation is employed to stabilize the variance of the response prior to an analysis of variance, thereby satisfying one of the assumptions implicit in the analysis of variance F tests. This method of analysis is satisfactory if one is solely interested in tests of hypotheses, but it is less than satisfactory if estimation is the goal. In many cases the experimenter will know in advance that the Poisson mean response must vary, at least an infinitesimal amount, as the levels of the independent variables are changed. Indeed, he might expect the change in mean response to be substantial. When this is the case, he wishes to estimate the difference in mean response for particular treatments (factor level combinations) or, when one or more of the independent variables are quantitative, he will carry the estimation process further and fit a response curve or surface.

Statistical analysis by transformation is unsuitable when estimation is the goal because one is estimating the expected value of the transformed data. Thus the square root transformation of a Poisson random variable, y, leads to the estimation of $E\{\sqrt{y}\}$. Squaring the least square estimate of $E\{\sqrt{y}\}$ does not lead to the "best" estimate of the Poisson mean, $E\{y\}$. We demonstrate this fact with a small experiment.

N = 1000 samples of size, n = 10, were collected on a Poisson random variable, y, with mean m = 2. Two confidence intervals were constructed for each sample. Interval Estimation Method #1 used the untransformed data with the intervals computed according to the formula,

$$\bar{y} \pm 2\sqrt{\bar{y}/10} \text{ , where } \bar{y} \text{ denotes the sample}$$

mean. The same sample response measurements were used for Interval Estimation Method #2 except that each observation was transformed to y* where y* = $\sqrt{y}$. Then the traditional least square confidence interval was computed using

$$\bar{y}* \pm 2s/\sqrt{10}$$

where

$$s = \sqrt{\sum_{i=1}^{10} \frac{(y_i^* - \bar{y}*)^2}{9}}$$

Then each endpoint of the interval was squared to obtain a confidence interval on E(y).

A comparison of the two interval estimation methods can be made by analysing the confidence intervals generated by the two methods for the 1000 samples of size n = 10. Table 1 shows the average and standard deviation of the width of the intervals for the two methods, the average value of the centerpoint and the fraction of times the intervals enclosed m.

The remainder of this article has been reproduced photographically from the authors manuscript.

Table 1: A Comparison of Confidence
Intervals for m Using Transformed and Untransformed Data

	Raw Data Intervals	Transformed Data Intervals
Average width	1.776	1.948
Standard Deviation of Widths	0.200	0.475
Fraction of Time m=2 is Enclosed by the Intervals	0.949	0.934
Average Midpoint of the Intervals	1.997	2.284

Note that Method #1 appears to be superior to the analysis
based on the transformed data. The interval widths possess a
smaller average and much less variation.

Most experimenters, using factorial experiments in trans-
formed Poisson data, employ an equal number of observations per
factor level combination. This allocation is far from optimal
for estimating $E(y)$, the response surface for the untransformed
data. Consequently, it is necessary to consider unequal
allocation of a sample to the factor level combinations of a
complete factorial experiment. We call this type of design a
"weighted factorial experiment."

This paper discusses the use of least squares in the analysi
of Poisson data when the data have been collected from a weighted
factorial experiment. Particularly, we give an expression for
$\hat{y}$, the estimator of $E(y)$, and give a closed form expression for
its variance. This permits us to choose optimal allocation
of a sample to the factor level combinations, or equivalently,
to select an optimal design.

2. Fitting Response Surfaces for Weighted Factorial Experiments

We will assume that the expected mean Poisson response,
$E(y)$, is related to a set of k independent variables, representi
the k factors in a weighted factorial experiment.

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_r x_r$$

We assume that the complete factorial contains s level combinations and that $\sum_{i=1}^{s} n_i = n$. The independent variables, $x_1, x_2, \ldots, x_r$, are orthogonal polynomials associated with the k factors contained in the experiment and y_i is the average of n_i observations taken at the ith factorial level combination. The form of the response surface, that is its order, is only limited by the number of levels included in the weighted factorial experiment.

Let,

$$E(\underline{Y}) = \underline{X}\underline{\beta} \tag{1}$$

where

$$\underline{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_r \end{bmatrix}$$

and

$$\underline{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1r} \\ 1 & x_{21} & x_{22} & \cdots & x_{2r} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{s1} & x_{s2} & \cdots & x_{sr} \end{bmatrix} = \begin{bmatrix} \underline{u}_1' \\ \underline{u}_2' \\ \vdots \\ \underline{u}_s' \end{bmatrix} \quad \text{and } r < s$$

Then it is well known that the least squares fit to equation (1) is

$$\hat{\underline{\beta}} = (\underline{X}'\underline{X})^{-1}\underline{X}'\underline{Y} \tag{2}$$

where

$$\underline{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_s \end{bmatrix}$$

The variance-covariance matrix of $\underline{Y}$ is the diagonal matrix,

$$V = \begin{bmatrix} v_1 & & \\ & v_2 & \\ & & \ddots \\ & & & v_s \end{bmatrix} \quad \text{where } v_i = \frac{E(y_i)}{n_i}, \quad i=1,2,\ldots,s.$$

The estimated mean response at a point, p, in the experimental region of the original k independent variables is obtained by substituting the values of these variables into the orthogonal polynomials, $x_1, x_2, \ldots, x_r$, to obtain $\hat{y} = \underline{a}'\hat{\underline{\beta}}$, where

$$\underline{a}' = [1 \quad x_{p1}, \quad x_{p2}, \quad \ldots, \quad x_{pr}].$$

Then

$$E(\hat{y}) = \underline{a}'\underline{\beta} \tag{3}$$

and

$$V(\hat{y}) = \underline{a}'(\underline{X}'X)^{-1}X'VX(X'X)^{-1}\underline{a}. \tag{4}$$

Letting

$$\underline{a}^{*'} = \underline{a}'(X'X)^{-1}$$

and

$$w_i = (\underline{a}^{*'}\underline{u}_i) \qquad i = 1, 2, \ldots, s \qquad (5)$$

we can rewrite $\hat{y}$ and $V(\hat{y})$ as

$$\hat{y} = \sum_{i=1}^{s} w_i y_i ,$$

$$V(\hat{y}) = \sum_{i=1}^{s} w_i^2 v_i . \qquad (6)$$

Our objective is to find expressions for the weights, w_i, $i=1,2,\ldots,s$, to obtain simple expressions for $\hat{y}$ and $V(\hat{y})$. We can then find the optimal allocation, $n_1, n_2, \ldots, n_s$ to the s factor level combinations to minimize $V(\hat{y})$ at a pre-selected point, p, in the experimental region. We will give the expressions for the weights for first order and complete factorial models for 2^k weighted factorial experiments and for second order and complete factorial models for 3^k weighted factorial experiments.

3. Weights for 2^k Weighted Factorial Experiments

Let $x_1, x_2, \ldots, x_k$ represent the k linear orthogonal polynomials corresponding to factors $1, 2, \ldots, k$. The complete 2^k

factorial model contains β_0 and terms corresponding to all products of $x_1, x_2, \ldots, x_k$ taking the variables one at a time, two at a time,...., and k at a time. Note that this model will contain exactly 2^k terms.

Then the $\underline{u}_i'$ vector associated with the ith factor level combination, $i=1,2,\ldots,s$, is

$$\underline{u}_i' = [1, e_{i1}, e_{i2}, \ldots, e_{ik}, e_{i1}e_{i2}, e_{i1}e_{i3}, \ldots, e_{i,k-1}e_{i,k}]$$

(7)

where

$$e_{ij} = \begin{cases} -1 & \text{if factor j is at the low level} \\ 1 & \text{if factor j is at the high level} \end{cases}$$

$$j=1,2,\ldots,k$$

Then it can be shown that the corresponding weight, w_i, $i=1,2,\ldots,s$, is

$$w_i = \underline{a}^{*'}\underline{u}_i = \frac{1}{2^k}\prod_{j=1}^{k}(1 + e_{ij}x_j)$$

(8)

Similarly, it can be shown that w_i, $i=1,2,\ldots,s$, for a first order model,

$$E(y) = \beta_0 + \sum_{i=1}^{k} \beta_i x_i,$$

is

$$w_i = \frac{1}{2^k}(1 + e_{i1}x_1 + e_{i2}x_2 + \ldots + e_{ik}x_k).$$

(9)

Thus for a weighted 2x2 factorial, the prediction equations for the complete factorial and first order models are, respectively,

$$\hat{y} = 1/4[(1-x_1)(1-x_2)y_1 + (1-x_1)(1+x_2)y_2$$
$$+ (1+x_1)(1-x_2)y_3 + (1+x_1)(1+x_2)y_4] \qquad (10)$$

and

$$\hat{y} = 1/4[(1-x_1-x_2)y_1 + (1-x_1+x_2)y_2$$
$$+ (1+x_1-x_2)y_3 + (1+x_1+x_2)y_4] \qquad (11)$$

where y_i is the mean response for n_i observations at the levels of x_1 and x_2 shown below:

i	x_1	x_2
1	-1	-1
2	-1	1
3	1	-1
4	1	1

4. Weights for 3^k Weighted Factorial Experiments

Let x_i, $i=1,2,\ldots,k$, represent the linear orthogonal polynomial corresponding to factor i and let x_{k+i} represent the second order orthogonal polynomial for the ith factor. Then a complete 3^k factorial model contains terms corresponding to x_j, $j=1,2,\ldots,2k$, plus terms corresponding to all combinations of products of these variables taking them two at a time, three at a time,...., k at a time, excluding terms that include products

-191-

of x_i and x_{k+i}, $i=1,2,\ldots$ or k. Note that this model will contain exactly 3^k terms. Thus for a complete 3^2 factorial,

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_1 x_2$$
$$+ \beta_6 x_1 x_4 + \beta_7 x_2 x_3 + \beta_8 x_3 x_4 \qquad (12)$$

where x_3 is a second order orthogonal polynomial in x_1 and x_4 is a similar second order orthogonal polynomial in x_2.

$$\text{Let } e_{ij} = \begin{cases} -1 & \text{if factor } j \text{ is at the low level} \\ 0 & \text{if factor } j \text{ is at the intermediate level} \\ 1 & \text{if factor } j \text{ is at the high level} \end{cases}$$
$$j=1,2,\ldots,k.$$

Then the $\underline{u}_i'$ vector associated with the ith factor level combination, $i=1,2,\ldots,s$, is

$$\underline{u}_i' = [1, e_{i1}, e_{i2}, \ldots, e_{ik}, (3e_{i1}^2-2), (3e_{i2}^2-2), \ldots, (3e_{ik}^2-2),$$
$$e_{i1}e_{i2}, e_{i1}e_{i3}, \ldots, (3e_{i1}^2-2)(3e_{i2}^2-2)\ldots(3e_{ik}^2-2)] \qquad (13)$$

(Note that $\underline{u}_i'$ contains 3^k elements in the same order as they appear in the complete factorial model.)

The second order model for a 3^k weighted factorial experiment is obtained by deleting all terms from the complete 3^k factorial model that are of a third or higher order. Deleting the corresponding elements from the $\underline{u}_i'$ vector for the complete

-192-

factorial model gives the $\underline{u}_i'$ vector for the second order model.

It can be shown that the weights, w_i, $i=1,2,\ldots,s$, $(s=3^k)$ for the complete factorial model are,

$$w_i = \underline{a}^{*'}\underline{u}_i = \prod_{j=1}^{k} h(e_{ij},x_j) \qquad (14)$$

where

$$h(e_{ij},x_j) = \frac{e_{ij}}{2} x_j(1+e_{ij}x_j) + (1-x_j^2)(1-e_{ij}^2).$$

It is difficult to obtain a simple expression for the weights, w_i, $i=1,2,\ldots,s$, for the second order model based on a weighted 3^k factorial experiment, but they can be obtained from the product,

$$w_i = \underline{a}^{*'}\underline{u}_i$$

where $\underline{u}_i' = [1, e_{i1}, e_{i2}, \ldots, e_{ik}, (3e_{i1}^2 - 2), \ldots, (3e_{ik}^2 - 2),$

$$e_{i1}e_{i2}, \ldots, e_{i,k-1}e_{i,k}]. \qquad (15)$$

For the 3^2 weighted factorial experiment we can give the prediction equation using the following three functions:

$$Q_1(x_1,x_2) = \frac{x_1 x_2}{4} (1 + x_1)(1 + x_2)$$

$$Q_2(x_1,x_2) = \frac{x_1}{2} (1 + x_1)(1 - x_2^2)$$

$$Q_3(x_1,x_2) = (1 - x_1^2)(1 - x_2^2).$$

Then

$$\hat{y} = Q_1(-x_1,-x_2)y_1 + Q_2(-x_1,x_2)y_2 + Q_1(-x_1,x_2)y_3$$

$$+ Q_2(-x_2,x_1)y_4 + Q_3(x_1,x_2)y_5 + Q_2(x_2,x_1)y_6$$

$$+ Q_1(x_1,-x_2)y_7 + Q_2(x_1,x_2)y_8 + Q_1(x_1,x_2)y_9 , \qquad (16)$$

where y_i, $i=1,2,\ldots,9$ is the mean response of n_i observations
taken at the following levels of x_1 and x_2:

i	x_1	x_2
1	-1	-1
2	-1	0
3	-1	1
4	0	-1
5	0	0
6	0	1
7	1	-1
8	1	0
9	1	1

Similarly by defining the functions

$$\Psi_1(x_1,x_2) = \frac{1}{6}\left[x_1^2 + x_2^2 + x_1 + x_2 + \frac{3x_1x_2}{2} - \frac{2}{3}\right] ,$$

$$\Psi_2(x_1,x_2) = \frac{1}{6}\left[x_1^2 - 2x_2^2 + x_1 + \frac{4}{3}\right],$$

and

$$\Psi_3(x_1,x_2) = \frac{1}{9}\left[5 - 3(x_1^2 + x_2^2)\right]$$

we can write the prediction equation for the second order model of the 3^2 weighted factorial experiment as

$$\hat{y} = \Psi_1(-x_1, -x_2)y_1 + \Psi_2(-x_1, x_2)y_2 + \Psi_1(-x_1, x_2)y_3$$

$$+ \Psi_2(-x_2, x_1)y_4 + \Psi_3(x_1, x_2)y_5 + \Psi_2(x_2, x_1)y_6$$

$$+ \Psi_1(x_1, -x_2)y_7 + \Psi_2(x_1, x_2)y_8 + \Psi_1(x_1, x_2)y_9 . \quad (17)$$

5. <u>Sample Allocation to Minimize $V(\hat{y})$ at a Point, p, in the Experimental Region</u>

We noted earlier that, for any weighted 2^k or 3^k factorial experiment,

$$V(\hat{y}) = \sum_{i=1}^{s} w_i^2 v_i \quad \text{where} \quad v_i = \frac{E(y_i)}{n_i}$$

and w_i, $i=1,2,\ldots,s$ are functions of $x_{p1}, x_{p2}, \ldots, x_{pk}$, the coordinates of a point in the experimental region where we want to estimate $E(y)$. We wish to find $n_1, n_2, \ldots, n_s$, the allocation of observations to the factor level combinations, that minimize $V(\hat{y})$ subject to the condition that $\sum_{i=1}^{s} n_i = n$. Assuming equal costs for all observations, the minimizing solution can be shown to be,

$$n_i = n_s \sqrt{\frac{E(y_i)}{E(y_s)} \frac{w_i^2}{w_s^2}} \quad (18)$$

and

$$n_s = \frac{n}{1 + \sum_{i=1}^{s-1} \sqrt{\frac{E(y_i)w_i^2}{E(y_s)w_s^2}}} . \quad (19)$$

-195-

How do you find the value of $E(y_i)$, $i=1,2,\ldots,s$, needed to make an optimal allocation of the n observations to the s factor level combinations? Since we will not know the true response surface, you can approximate these values by making a trial run at the s factor level combinations. A preliminary computation of $\hat{y}$ will provide the approximating values for $E(y_i)$, $i=1,2,\ldots,s$.

6. Comments

Most response surface explorations utilize first or second order linear models. Factors are held at two levels when fitting a first order model, and three levels for a second order model. The weights and optimal sample allocations for these situations were presented in Sections 3, 4, and 5. Although of lesser importance, weights for a 4^k, 5^k, or, in general b^k, $b=2,3,4,\ldots$, can be obtained using the procedure described in Section 2.

The variance of $\hat{y}$ for a particular estimation problem is obtained by approximating $v_i = \dfrac{E(y_i)}{n_i}$ with $\dfrac{y_i}{n_i}$, $i=1,2,\ldots,s$. Note that this approximation will be quite good because y_i is the average of n_i observations at the ith factor level combination.

An approximate bound on the error of estimation for $E(y)$ is then $2\sqrt{\hat{V}(\hat{y})}$ where $\hat{V}(\hat{y})$ is the value of $V(\hat{y})$ with $E(y_i)$ replaced by y_i, $i=1,2,\ldots,s$. Finally, note that $V(\hat{y})$ will vary depending upon the location of the point, p, where you wish to estimate $E(y)$. The experimenter should decide where maximum information on $E(y)$ is desired prior to making the sample allocation.

INFERENCES ON FUNCTIONS OF THE PARAMETERS

OF UNIVARIATE DISTRIBUTIONS

Ronald L. Racicot

Applied Mathematics and Mechanics Division
Benet Weapons Laboratory
US Army Watervliet Arsenal
Watervliet, New York 12189

1. INTRODUCTION

The basic problem considered is the following: given the analytical form of a univariate probability distribution function defined by one or more unknown parameters and given a sample from the distribution, determine statistical inference information on a function of the unknown parameters. Of particular interest are the more difficult problems involving small samples for which the accuracy of asymptotic solutions is doubtful and for quantities of interest which are functions of more than one population parameter. No general solution exists to this problem for the classical frequency confidence intervals but solution for a Bayesian interval can often be found. The main difficulty in the Bayesian approach is, of course, making a suitable choice of the prior distribution on the population parameters or quantity of interest.

The remainder of this article was reproduced photographically from the manuscript supplied by the author.

The approach to this problem discussed in this paper is to employ a Bayesian technique which is, for a particular class of probability distributions to be defined later, computationally more efficient than the usual Bayesian techniques based on the likelihood function. This then facilitates computer studies of the exactness of Bayesian confidence intervals from the classical frequency viewpoint and in studies for the possible development of classical frequency intervals from the Bayesian distributions.

The Bayesian technique employed makes use of the conditional probability distributions of the estimators of population parameters given the true values of the parameters. This is in contrast to the use of the likelihood function which gives the distribution of the sample data directly. There is consequently, at the outset, a general difference in the two Bayesian confidencing approaches. The confidence intervals derived from the estimator distribution, although exact in a Bayesian sense, depend on the particular estimator used. Also, if the parameter estimators are not sufficient statistics then this gives additional cause for differences in comparison to the likelihood approach. In the work to be discussed in this paper only distributions of the maximum likelihood estimators were studied because of the desirable properties of these estimators.

Two examples of the general problems of interest are given as follows:

a. For the Weibull distribution $f(x;\alpha,\beta)$ is given as

$$f(x;\alpha,\beta) = (\frac{\beta}{\alpha}) (\frac{x}{\alpha})^{\beta-1} e^{-(x/\alpha)^\beta}$$

(1)

The mean μ is then given by

$$\mu = \alpha\Gamma \ (1 + 1/\beta) = g \ (\alpha,\beta) \tag{2}$$

where confidence intervals are required for the mean.

b. For arbitrary first failure distribution $F(t;\alpha,\beta)$ the "ideal repair" interval reliability at arbitrary time t for interval τ is given by the expression

$$R(t,\tau;\alpha,\beta) = 1 - F(t + \tau;\alpha,\beta)$$

$$+ \int_0^t [1 - F(t + \tau - x;\alpha,\beta)]h(x;\alpha,\beta)dx \tag{3}$$

where $h(x;\alpha,\beta)$ is determined from the renewal equation

$$h(x;\alpha,\beta) = f(x;\alpha,\beta) + \int_0^x f(x - y;\alpha,\beta) \ h(y;\alpha,\beta)dy \tag{4}$$

As can be seen the interval reliability for given t and τ is a function of the unknown distribution parameters α and β. Confidence intervals are then required for the interval reliability.

2. ONE PARAMETER DISTRIBUTION

The case of the one parameter distribution is relatively straightforward and generally offers no difficulty. It is used here to introduce some of the basic ideas in the construction of confidence intervals. For a given sample of size n from the distribution $F(x;\alpha)$ various techniques can be used to determine an estimate $\hat{\alpha}$ of the parameter α. Maximum likelihood is basically the estimating technique used in the work described in this paper. It is also a technique that often leads to minimum variance for estimates and has other desirable properties (Reference 1, Section 33.2).

One approach to determining classical frequency confidence intervals for either α or a suitably behaved function $g(\alpha)$ is to generate the conditional probability distribution of $\hat{\alpha}$ (or $g(\hat{\alpha})$) as a function of α (or $g(\alpha)$); that is $P_{\hat{\alpha}}(\hat{\alpha};\alpha)$ for the estimator $\hat{\alpha}$. One then finds limits $\gamma_1(\alpha)$ and $\gamma_2(\alpha)$ for $\hat{\alpha}$ as a function of α such that the probability that $\hat{\alpha}$ falls between γ_1 and γ_2 is fixed at the desired percent confidence level p. The limits $\gamma_1(\alpha)$ and $\gamma_2(\alpha)$ form curves on the $(\hat{\alpha},\alpha)$ plane with α being the independent variable. One can now go through various arguments to show that for a given estimator $\hat{\alpha}$, the values of α for which $\hat{\alpha}=\gamma_1(\alpha_1)$ and $\hat{\alpha}=\gamma_2(\alpha_2)$ are the upper and lower confidence limits on α (Reference 1, Section 34.2). With confidence intervals constructed in this manner, independent trials, where each trial consists of drawing say a sample n from the given population, would yield confidence limits that covered the true value an average of p percent of the trials. In this case the true parameter can be regarded as fixed or allowed to vary in an arbitrary fashion.

In the Bayesian approach the parameter α is considered itself to be a random variable having a certain a priori distribution $f_1(\alpha)$. One then uses Bayes' theorem to determine the a posteriori distribution $f_2(\alpha)$ of α for a given sample outcome making use of the conditional distribution of the sample as a function of α. The conditional distribution of the sample for given α is often assumed to be the likelihood function for the more complicated problems. The conditional distribution of an estimator $\hat{\alpha}$ for given α can also be used to generate the a posteriori distribution of α.

Bayes' theorem for both of these cases are given as follows:

$$f_2(\alpha/\overline{x}) = C_1 \, f_1(\alpha) \, L \, (\overline{x};\alpha) \tag{5}$$

and

$$f_2(\alpha/\hat{\alpha}) = C_2 \, f_1(\alpha) \, f \, (\hat{\alpha};\alpha) \tag{6}$$

in which L is the likelihood function, $\overline{x}$ is the sample outcome, $\hat{\alpha}$ is the estimator of α and C_1 and C_2 are normalizing constants. Confidence intervals can then be constructed directly from the distribution $f_2 \, (\alpha/\overline{x})$ or $f_2 \, (\alpha/\hat{\alpha})$.

The main difficulty in the Bayesian approach is to determine the a priori distribution $f_1 \, (\alpha)$, particularly when there is very little and/or questionable prior data on the true parameter values one could possibly expect.

The likelihood function is most often used in constructing Bayesian confidence limits because there is no necessity for deciding what estimator to use and because of the general ease in theoretically constructing a solution for confidence limits. Also, determining the two-dimensional function $f \, (\hat{\alpha};\alpha)$ can be tedious if analytical solutions are not possible. As will be shown shortly, however, the quantities $\hat{\alpha}$ and α can often be transformed into a single parameter $\hat{\alpha}_s$, the probability distribution of which is independent of the true parameter α (eg. $\hat{\alpha}_s = \hat{\alpha}/\alpha$ for the exponential distribution). The problem is thus reduced to the determination of a one-dimensional function $f \, (\hat{\alpha}_s)$ for given sample size n.

This type of transformation becomes much more important in the multi-parameter problem. Also, note that the distribution of $\hat{\alpha}_s$ yields directly the classical frequency interval for α which, in this instance, is equal to the Bayesian interval determined from $f_2(\alpha;\hat{\alpha})$ with uniform prior on α.

The primary reason for approaching the Bayesian confidencing problem using the estimator distribution is that once the distribution of the estimator is available for given sample size n, the computer time for constructing confidence intervals was found to be much shorter than when using the likelihood approach for the more complicated problems of interest. This then facilitated computer studies of the exactness, from a classical frequency viewpoint, of the Bayesian intervals and in studies of procedures for generating classical frequency intervals from the Bayesian probability distributions.

3. TWO OR MORE PARAMETER DISTRIBUTIONS

Only the two parameter problem will be considered in what follows to simplify the discussion although extension to more than two parameters is straightforward. In many instances, however, only the two parameter problem is practical to solve.

When confidence intervals are desired for a function g of two unknown population parameters, the classical frequency approach discussed in the previous section cannot be applied in general. For example, the distribution of the estimator of say $\hat{g} = g(\hat{\alpha},\hat{\beta})$ does not generally depend directly on the true value of g but rather on both of the unknown parameters α and β.

Hence, for a fixed true value of g, α and β can take on different values yielding different distributions of $\hat{g}$. There are exceptions to this such as the first failure reliability for the two parameter Weibull distribution (Reference 4).

In the Bayesian approach the key to solving the two parameter problem is the determination of the bivariate distribution $f(\hat{\alpha},\hat{\beta};\alpha,\beta)$ for the estimators $\hat{\alpha}$ and $\hat{\beta}$. Once the bivariate distribution is available for $\hat{\alpha}$ and $\hat{\beta}$ as a function of α and β and by assuming some prior distribution on the unknown parameters the distribution can be transformed as in the one parameter case into a posteriori distribution of α and β for given $\hat{\alpha}$ and $\hat{\beta}$. The distribution of a function $g(\alpha,\beta)$ of the parameters α and β can then be determined from $f(\alpha,\beta;\hat{\alpha},\hat{\beta})$ (Reference 2, Chapters 5 and 7). The determination of exact Bayesian confidence intervals can then be obtained as in the one parameter case.

The function $f(\hat{\alpha},\hat{\beta};\alpha,\beta)$ is a four-dimensional function. In the next section the transformation of the four-dimensional into a two-dimensional problem is considered. This then greatly simplifies the computational aspects of this overall problem.

4. A GENERAL TRANSFORMATION ON ESTIMATORS:

Consider the maximum likelihood estimators of the parameters α and β for a given sample x_i, $i = 1, 2, \ldots, n$ taken from a population with distribution $f(x;\alpha,\beta)$. The likelihood function L in this case is given by

$$L = \prod_{i=1}^{n} f(x_i;\hat{\alpha},\hat{\beta})$$

(7)

-203-

where the estimators $\hat{\alpha}$ and $\hat{\beta}$ maximize the function L. The solution for $\hat{\alpha}$ and $\hat{\beta}$ in this case depend only on the sample values x_i and the functional form $f(x_i;\hat{\alpha},\hat{\beta})$ or alternately on the functional form of the cumulative distribution function $F(x_i;\hat{\alpha},\hat{\beta})$. Estimators other than maximum likelihood, of course, could fit this criteria.

Consider next the original cumulative probability distribution $F(x;\alpha,\beta)$ for some given true α and β and from which the sample x_i is generated. It can be shown that for an arbitrary sample value x_i the functional value of $F(x_i;\alpha,\beta)$ evaluated at the point x_i is itself a random variable, call it R_i, which is distributed uniformly on the interval (0, 1) Reference 3, Page 313). In this case then $R_i = F(x_i;\alpha,\beta)$ which is independent of α and β. Using this expression to solve for x_i gives

$$x_i = F^{-1}(R_i;\alpha,\beta).$$

(8)

Substituting this equation into the function $F(x_i;\hat{\alpha},\hat{\beta})$ used to determine the estimators $\hat{\alpha}$ and $\hat{\beta}$ finally gives

$$F(x_i;\hat{\alpha},\hat{\beta}) = F(F^{-1}(R_i;\alpha,\beta);\hat{\alpha},\hat{\beta})$$

(9)

If this function can be transformed into the function $F(Z(R_i);\hat{\alpha}_s,\hat{\beta}_s)$ which has the same analytical form as $F(x_i;\hat{\alpha},\hat{\beta})$, then it is clear that the estimators $\hat{\alpha}_s$ and $\hat{\beta}_s$ (or any function of them) in this expression are statistically independent of α and β since the $Z(R_i)$ are independent of α and β. In general $\hat{\alpha}_s = u(\alpha,\beta,\hat{\alpha},\hat{\beta})$ and $\hat{\beta}_s = v(\alpha,\beta,\hat{\alpha},\hat{\beta})$. This means that it is only necessary to generate the joint distribution of $\hat{\alpha}_s$ and $\hat{\beta}_s$ to completely define the conditional frequency distribution $f(\hat{\alpha},\hat{\beta};\alpha,\beta)$.

Thus the four-dimensional problem is reduced to a two-dimensional problem. The key to this problem is transforming $F(F^{-1}(R_i;\alpha,\beta);\hat{\alpha},\hat{\beta})$ into the form $F(Z(R_i);\hat{\alpha}_s,\hat{\beta}_s)$.

5. EXAMPLES OF THE GENERAL TRANSFORMATION

a. Uniform Distribution.

The cumulative distribution function for this case is given by

$$F(x;\alpha,\beta) = \begin{cases} 0 & ; \ x < \alpha \\ (x-\alpha)/(\beta-\alpha); & a \leq x < \beta \\ 1 & ; \ x \geq \beta \end{cases}$$

(10)

Letting $R_i = F(x_i;\alpha,\beta)$ and solving for x_i gives

$$x_i = R_i (\beta-\alpha) + \alpha$$

(11)

over the interval of interest

Substituting into the expression $F(x_i;\hat{\alpha},\hat{\beta})$ yields the function

$$F(Z_i(R_i);\hat{\alpha}_s,\hat{\beta}_s) = (Z_i-\hat{\alpha}_s)/(\hat{\beta}_s-\hat{\alpha}_s)$$

(12)

where $Z_i(R_i) = R_i$;

$\hat{\alpha}_s = (\hat{\alpha}-\alpha)/(\beta-\alpha)$;

$\hat{\beta}_s = (\hat{\beta}-\alpha)/(\beta-\alpha)$.

It is only necessary then to generate the joint distribution of $\hat{\alpha}_s$ and $\hat{\beta}_s$ to completely define the distribution of $\hat{\alpha}$ and $\hat{\beta}$ for given α and β.

b. Gumbel Extremal Distribution of Type I.

In this case

$$P(x) = 1 - \exp(-\exp[(X-\alpha)/\beta])$$ (13)

Using the previously described procedures it can be shown that

$$Z(R_i) = \ln[-\ln(1-R_i)];$$

$$\hat{\alpha}_s = (\frac{\hat{\alpha}-\alpha}{\beta});$$

$$\hat{\beta}_s = (\hat{\beta}/\beta).$$ (14)

c. Weibull Distribution.

For the two parameter Weibull distribution

$$P(x) = 1 - \exp\{-(\frac{X}{\alpha})^\beta\}$$ (15)

from which is obtained

$$Z(R_i) = -\ln(1-R_i);$$

$$\hat{\alpha}_s = (\hat{\alpha}/\alpha)^{\hat{\beta}};$$

$$\hat{\beta}_s = (\hat{\beta}/\beta).$$ (16)

The results given by the above expression was derived by Thoman, et al (Reference 4) in their work on inferences for the parameters of the Weibull distribution.

6. INFERENCES ON FUNCTIONS OF THE WEIBULL PARAMETERS

The two parameter Weibull distribution proves an interesting example for application of the statistical theory just presented. First, general analytical solutions for the distribution of the maximum likelihood

estimators are not available. The bivariate distribution of $\hat{\alpha}_s$ and $\hat{\beta}_s$ in this instance must therefore be determined numerically. This was done using Monte Carlo simulation for different sample sizes. Specifically the frequency distribution $f(x,y)$ where $x = (\hat{\alpha}/\alpha)^\beta$ and $y = \hat{\beta}/\beta$ was generated for sample sizes n = 5, 8 and 20 using 20,000, 20,000 and 10,000 points respectively. Other sample sizes could of course have been considered. Once the distribution $f(x,y)$ is generated for a given sample size, Bayesian confidence intervals on functions of the Weibull parameters can be computed.

A particular function considered for application was the mean $\mu = \alpha\Gamma(1 + 1/\beta)$. For given maximum likelihood estimates of the parameters, $\hat{\alpha}$ and $\hat{\beta}$, and assuming a uniform prior on α and β, the probability distribution of the mean from which confidence intervals can be computed is given as follows:

$$F_\mu(Z) = \text{Prob } [\mu \leq Z]$$

$$= \int_0^\infty \int_\gamma^\infty f(x, y)dxdy$$

$$\text{where} \quad = [\frac{\hat{\alpha}\Gamma(1+y/\hat{\beta})}{Z}]^\beta. \tag{17}$$

For example, an 80% Bayesian confidence interval can be determined by solving for Z_1 and Z_u such that $F_\mu(Z_1) = 0.1$ and $F_\mu(Z_u) = 0.9$. In this case Z_1 and Z_u are the lower and upper confidence limits respectively.

7. RESULTS OF COMPUTATIONS

A number of computations were performed using the above described procedures for confidencing the Weibull mean. Some of the results of these computations are described as follows:

a. **Exactness of Confidence Intervals.**

Monte Carlo simulation was used to generate sets of 5000 and 1000 samples of sizes 5 and 8 from a Weibull distribution with fixed true mean μ equal to 1.0 and a uniform prior on the shape parameter β. For each given sample, the function $F_\mu(Z)$ in equation (17) was evaluated for $Z = 1.0$. If the Bayesian confidence intervals determined using $F_\mu(Z)$ are exact then the computation of this function at $Z = 1$ for the given samples should yield a uniform distribution on *the interval* [0, 1]. This interval was divided into ten parts and a chi square test was performed on the resulting data. The hypothesis that the distribution of $F_\mu(Z = 1)$ was uniform could not be rejected down to the 50% confidence level for the cases considered.

The Bayesian intervals were also studied to determine exactness from a classical frequency viewpoint. In this instance the shape parameter β was held fixed as well as the mean μ. A number of Monte Carlo simulations were performed with β fixed at various values. For values of β greater than about 7.0 the Bayesian intervals were very nearly exact. For smaller values of β, particularly near 1.0, however, the intervals deviated from being exact although not by a great degree.

A number of approaches were taken to attempt to transform the Bayesian intervals into exact classical frequency intervals for all values of β. The most successful approach studied thus far involves the use of a biasing factor ζ on the true mean which is a function of the unbiased maximum likelihood estimator for β. In the Monte Carlo trials for exactness described above, instead of evaluating $F_\mu(Z = \mu_t)$ where μ_t, the true mean, is fixed at 1.0, the function $F_\mu(Z = \mu_t \zeta)$ is evaluated and exactness is then checked for this function. For the Weibull mean, the function

$$\zeta = 1.0 + \frac{a}{\hat{\beta}^*} + \frac{b}{(\hat{\beta}^*)^2} \, , \tag{18}$$

in which $\hat{\beta}^*$ is the unbiased estimator for β, gave confidence intervals that were very nearly exact for fixed values of β ranging from 1.0 to 10.0. For a sample size of 8, for example, the empirically derived values of a and b required in equation (18) were approximately 0.055 and 0.015 respectively. In this case values of ζ equal to 1.07, 1.0116 and 1.0056 are obtained for $\hat{\beta}^*$ equal to 1.0, 5.0 and 10.0 respectively. The Bayesian confidence limits when divided by ζ yield confidence limits which are nearly exact from a classical frequency viewpoint. No general conclusion will be made at this time on this overall approach since work on this problem is not complete. Further studies are to be conducted in the future.

b. Comparison of Confidence Limits.

Bayesian confidence intervals using uniform priors were determined on Monte Carlo sample data using both the likelihood function and the estimator distribution. The samples were generated assuming true parameter values $\mu = 1$ and $\beta = 1$ and 3. Tables 1 and 2 present some of the results of these computations for a sample size of 5.

No major conclusions can be made based on the data generated thus far. It does appear, however, that the two Bayesian confidence limits are similar in their general behavior although some differences are observed in their values depending on the value of β as can be seen from the data in Tables 1 and 2.

Another general observation is that the computer times required to generate the Bayesian intervals based on estimator distribution, once the bivariate distribution of parameter estimators is generated, were considerably less than for the Bayesian intervals based on the likelihood function. An order of magnitude difference was observed for many of the confidencing problems considered.

Table 3 lists some results comparing the Bayesian and classical frequency lower 90% confidence limit for a sample size of 8. As indicated previously the classical frequency intervals for the values of β considered agree more closely for the larger values of β.

c. Moments of the Maximum Likelihood Mean.

It is of interest to observe the asymptotic behavior of estimators as a function of sample size. It is known for example that maximum likelihood estimators are asymptotically normal. Table 4 lists the resulting computation of the variance, skewness and kurtosis of the mean as a function

-210-

of β and sample size. One property of the normal distribution is that it has a skewness of zero and kurtosis equal to 3.0. As can be seen for $\beta > 1$ the maximum likelihood estimators for the mean have approached the normal distribution even for the sample size of 5.

8. REFERENCES:

(1) Cramer, H., Mathematical Methods of Statistics, Princeton University Press, Princeton, 1963.

(2) Papoulis, A., Probability, Random Variables, and Stochastic Processes, McGraw Hill, New York 1965.

(3) Parzen, E., Modern Probability Theory and Its Application, Wiley, New York, 1960

(4) Thoman, D. R., Bain, L. J., and Antle, C. E., "Inferences on the Parameters of the Weibull Distribution," Technometrics, Vol. 11, No. 3, August 1969, pp. 445-460.

TABLE 1: TWO-SIDED BAYESIAN CONFIDENCE INTERVALS ON THE

WEIBULL MEAN FOR $\beta = 1.0$: CONFIDENCE LEVEL = 80%

$\mu = 1.0$ AND SAMPLE SIZE = 5.

CONDITIONAL DISTRIBUTION USED

SAMPLE NUMBER	MAXIMUM LIKELIHOOD ESTIMATOR DISTRIBUTION		LIKELIHOOD FUNCTION	
	LOWER	UPPER	LOWER	UPPER
1	.577	1.103	.638	1.062
2	.623	1.502	.699	1.399
3	.320	2.561	.414	2.430
4	.635	1.136	.687	1.097
5	.721	3.442	.874	3.059
6	.234	.743	.271	.674
7	.886	4.580	1.083	3.888
8	.402	1.202	.463	1.078
9	.379	1.712	.455	1.461
10	.502	1.285	.565	1.125

TABLE 2: TWO-SIDED BAYESIAN CONFIDENCE INTERVALS ON THE

WEIBULL MEAN FOR β = 3.0: CONFIDENCE LEVEL = 80%,

μ = 1.0 AND SAMPLE SIZE = 5.

CONDITIONAL DISTRIBUTION USED

SAMPLE NUMBER	MAXIMUM LIKELIHOOD ESTIMATOR DISTRIBUTION		LIKELIHOOD FUNCTION	
	LOWER	UPPER	LOWER	UPPER
1	.660	.976	.704	.983
2	.810	.985	.836	.987
3	.968	1.618	1.054	1.636
4	.842	1.281	.902	1.291
5	.602	1.403	.685	1.462
6	.994	1.266	1.036	1.269
7	.951	1.317	1.004	1.327
8	.992	1.560	1.067	1.573
9	.806	1.288	.871	1.306
10	.888	1.466	.961	1.480

TABLE 3: BAYESIAN AND CLASSICAL LOWER 90% CONFIDENCE LIMITS
ON THE WEIBULL MEAN FOR SAMPLE SIZE OF 8 AND TRUE
MEAN OF 1.0.

SAMPLE NUMBER	TRUE β = 1.0		TRUE β = 3.0		TRUE β = 5.0	
	BAYESIAN	CLASSICAL	BAYESIAN	CLASSICAL	BAYESIAN	CLASSICAL
1	.562	.503	.737	.715	.809	.795
2	.785	.735	.910	.893	.939	.928
3	.902	.808	.864	.838	.891	.875
4	.922	.858	.947	.927	.959	.947
5	.562	.538	.858	.847	.916	.909
6	1.243	1.136	1.005	.980	.985	.970
7	.538	.516	.848	.837	.910	.903
8	.841	.767	.879	.856	.908	.894
9	.686	.612	.782	.758	.837	.822
10	.639	.594	.836	.818	.889	.878

TABLE 4: MOMENTS OF THE MAXIMUM LIKELIHOOD MEAN

B	SAMPLE SIZE = 5			SAMPLE SIZE = 8			SAMPLE SIZE = 20		
	VARIANCE	SKEWNESS	KURTOSIS	VARIANCE	SKEWNESS	KURTOSIS	VARIANCE	SKEWNESS	KURTOSIS
0.5	1.917	8.416	194.55	.916	4.259	5.29	.263	1.415	6.50
1.0	.194	.799	3.78	.125	.701	3.69	.050	.464	3.37
1.5	.089	.383	2.94	.057	.361	3.09	.023	.254	3.13
2.0	.053	.198	2.78	.034	.207	2.96	.014	.155	3.07
2.5	.036	.084	2.75	.023	.113	2.93	.009	.093	3.04
3.0	.026	.005	2.75	.016	.047	2.92	.007	.050	3.03
3.5	.020	-.055	2.77	.012	-.002	2.92	.005	.018	3.03
4.0	.015	-.103	2.79	.010	-.042	2.92	.004	-.008	3.02
4.5	.012	-.141	2.81	.008	-.073	2.93	.003	-.029	3.02
5.0	.010	-.173	2.83	.007	-.099	2.94	.003	-.046	3.02

DYNOSS - DYNAMICALLY OPTIMIZED SMOOTHING SPAN

Roberto Fierro
Analysis and Computation Division
White Sands Missile Range, New Mexico

FOREWORD.

The work covered by this report was done in the Software Section, Data Reduction Branch, Analysis and Computation Division, White Sands Missile Range. This study is part of a continuing program to determine optimum methods for the reduction of missile data at White Sands Missile Range. The general subject of filtering/smoothing is very old in the field of data reduction, though the bulk of the work reported here was accomplished during the year 1970.

ABSTRACT.

This report describes a dynamic recursive least squares polynomial smoothing span technique. Details of theory, design, and operation of the technique are given. This method is compared with classical least squares fixed span smoothing. The technique is not limited to any given type of data but is amenable to time series measurements from any given source. Typical measurements come from radar, fixed camera, or tracking camera. The computing time is extremely small and invariant with length of span. DYNOSS is being applied to missile trajectory data for some current testing programs at White Sands Missile Range. By every important criterion, DYNOSS is always better than the least squares fixed span smoothing method.

The remainder of this article was reproduced photographically from the authors manuscript.

I. INTRODUCTION.

The problems of smoothing/filtering data are old in the field of data reduction. For a good number of years, the fixed span least squares smoothing method has been satisfactory and no significantly new methods or approaches have been found necessary. However, the development of extremely high accelerated multi-stage missiles makes necessary a re-examination of past smoothing/filtering techniques. The purpose of the present paper is to show that the dynamic span smoothing approach offers significant advantages over fixed point smoothing in the generation of missile state estimates. Furthermore, the theory introduced is applicable to a wide variety of estimation problems. With the advantages that DYNOSS offers, it cannot be denied recognition as a significant advancement in the state of the art.

In this report, we refer to filters and filtering; however, the point being evaluated is the smoothed point. The filtered data (real time point) can be obtained by propagating the smoothed results to the end point.

To compare the performance of fixed span filters and DYNOSS, error free trajectory data was filtered over a fixed span and then through DYNOSS. The trajectories used to exercise the filters were a Standard Athena trajectory and an extremely high accelerated simulated trajectory. The Athena data consists of a first stage burnout, second stage ignition and burnout and the beginning of a ballistic trajectory. The extremely high accelerated (up to 10,000 ft/sec^2)

launch consists of short ignition periods, and abrupt changes in acceleration. The standard position, velocity and acceleration profile of the Z component of the trajectories are shown in Figs. 1, 2, 3, 14, 15, and 16. The Athena data was generated and used at 20 samples per second while the High Acceleration data was at 50 samples per second. Since, recovery of the true signal is also a function of noise content, random error was then added to the standard position data and the filtering repeated.

The results of experiments to compare the DYNOSS technique to that of fixed least squares smoothing are described in detail with graphs.

The author wishes to express his appreciation to Darold Comstock for his outstanding contributions to this study.

II. DYNOSS

A. Historical Background

The development of DYNOSS (acronym for Dynamically Optimized Smoothing Spans) came as a result of studying the specific control conditions that could be applied to the equations for the weighted least squares recursive estimation (WLSRE) filter.
The WLSRE filter/smoother can be described by the equations:

$$X = \hat{X} + (G + H^t WH)^{-1} H^t W (Z - H\hat{X}),$$

$$\sigma^2 = (G + H^t WH)^{-1},$$

$$\hat{X}_{t+1} = \phi X_t,$$

$$G_{t+1} = (\phi \sigma_t^2 \phi^t)^{-1},$$

where:

X = optimized state vector,

$\hat{X}$ = predicted state vector,

G = Weight matrix for $\hat{X}$,

W = Weight matrix for Z,

Z = observations,

σ^2 = variance-covariance matrix for X,

ϕ = transition matrix,

and H is the observation matrix which relates the observations to the state estimates.

Thus we see that a previous estimate may be updated by the amount that a new observation contributes.

The major problem areas in using the WLSRE equations are in determining W and in imposing checks and controls when the transition matrix does not adequately describe the process.

In the early stages of our study the comparison of classical least squares to desk calculated recursive estimations provided us with further insight into the WLSRE filter and aided in specifying the control conditions and equations to be used. It was at this point that we realised there existed a dire need of a technique for determining the optimum filtering span.

Data to be smoothed by polynomial filters are sampled at a high enough rate to conform to the filter as long as the number of point smoothing (K) is not too large. Regardless of the number of observations, the optimized filtered data will be obtained using the largest set (span) of input data that conforms to the filter design. Therefore it is highly desirable for each point to find the largest set of data which includes that point and conforms to the filter specification.

If the recursive equation is used as stated, then the span will be from the first data point through the latest data point. Thus it becomes necessary to find some method of restricting the span. Of the possible ways of restricting K (Number of data points used in the filter), there are three which we believe should be mentioned. These three are thrusting errors, fixed K and DYNOSS.

The true model usually consists of many more variables than

those listed in X. Therefore, the predicted state will be slightly
in error, since these variables were not considered when X at t was
transformed to $\hat{X}$ at $t + 1$. All these variables usually cannot be
extracted from the data nor obtained from other sources with the
desired degree of accuracy. However, one way in which they can be
taken into consideration is to include a thrusting error in G, the
weight matrix of the predicted state. This effectively reduces the
filtering span by assigning a smaller weight to the predicted state,
or it may be viewed as a relatively larger weight being assigned to
the observation. In using thrusting errors, it is usually necessary
to obtain the errors by monitoring accelerations, altitude, or some
other missile parameter. The main pitfall with this approach is
that we really do not include the thrusting force per se in the
equation of state used to generate the state estimates. Although
the augmentation of the G matrix with a thrusting error does enhance
the value of the state estimates. Another drawback is that usually
the checks made are for a given standard missile system. However,
we do not know what checks to provide for the missile systems of
the future or the missile that does not conform to the standard.

The weight matrix of the predicted state, G, may be varied
such that one effectively uses a fixed K. The main drawback with
using a fixed K is that one will obtain a misfit of the data during
such events as ignition and burnout. The ideal situation would be
to vary the weight matrix G such that K will be dynamically adjusted

-222-

to yield optimum results. This may also be viewed as selecting the optimum data span or set size.

DYNOSS does not exhibit the weaknesses of the other two methods. In fact DYNOSS can detect all events with remarkable accuracy and optimize the data through a dynamic process which selects the appropriate K. Furthermore, DYNOSS will handle any present and future missile system without any modifications and/or human intervention.

In utilizing the WLSRE equations, we need a procedure to compute, W, the weight matrix for Z. The possible ways to compute W are the following:

    (1) Compute W based on the difference between the observation and a predicted value of the observation.

    (2) Compute W from $W = \dfrac{1}{\sigma_Z^2}$ and $G = f\,(K,\ \sigma_Z^{\,2})$.

Computing W by the first method has certain disadvantages. Since W is based on the difference between predicted value of the observation and the observation, the error in the weight becomes a function of the error in both values. Thus an error in the predicted value results in an erroneous weight. This becomes critical, especially near event times, when the predicted value cannot reflect the event and as a result the latest observation is assigned a false weight. If this process is continued over several points, then the event may be removed from the filtered data.

-223-

The second approach has the disadvantage that all observations
are given the same weight. Further investigation reveals that this
is not very much of a disadvantage, since historically it has been
observed that the filtered output of a set of data with random errors
is not significantly degraded if the weights are chosen as being
equal, such as in averaging a set of numbers. The key point being
that the errors are random. This is the case for the types of data
being considered. This approach has the advantage that σ_z^2 can be
factored from G and W, leaving $(G + H^t WH)^{-1}$ as a function of K and Δt.
Thus the $(G + H^t WH)^{-1}$ matrix does not have to be inverted every time
but can be predicted based on the known values K and Δt.

B. Introduction

The purpose of DYNOSS is to optimize data through a dynamic
process which will increase or decrease the span size depending on
the outcome of a 99% confidence level F-test. Another extremely
attractive feature is the use of the WLSRE technique. The program
estimates the variance (σ_v^2) of the input data based on third order
ordinary differences and compares it to the variance (σ_c^2) of the
input data based on deviations from the model. Based on this test
the span size will be increased or decreased. The span size can
also be controlled externally if the user specifies the maximum
size of the span that he desires. When the program has increased
the span size such that it has reached the specified maximum, then

-224-

it will continue at this constant span size until the model no
longer fits the data and the program forces the set size to decrease.

C. Mathematical Formulation

Initially we determine the state estimates X, $\dot{X}$, and $\ddot{X}$ from least squares equation

$$X = (H^t H)^{-1} H^t Z. \tag{1}$$

Using X from Eq. (1) and, ϕ, the transition matrix, we see that

$$\hat{X}_{t+1} = \phi X_t. \tag{2}$$

Now we correct the prediction, $\hat{X}$, with the correction

$$C = (G + H^t W H)^{-1} H^t W (Z - H\hat{X}). \tag{3}$$

Eq. (3) reduces to

$$C = P(K, \Delta t) \ H^t \ (Z - H\hat{X}) \tag{4}$$

where P depends only on K (span size) and Δt (increment of time between observations).

The detailed computation of, Eq. (4) is considered in the Appendix. The corrected state becomes

$$X = \hat{X} + C.$$

To obtain the random noise content of the observations, use the equation

$$\sigma_V^2 = \frac{m!}{(2m)!} \frac{\Sigma(\delta_i^3)^2}{K-m} \qquad \begin{array}{l} K \geq 9 \\ i = t-\left(\frac{K-9}{2}\right), \; t-\left(\frac{K-11}{2}\right), \ldots, \; t+\left(\frac{K+1}{2}\right) \end{array} \qquad (5)$$

where

$$\delta_i^3 = Z_i - 3Z_{i-1} + 3Z_{i-2} - Z_{i-3},$$

σ_V^2 = variance of observations,

m = order of differences (3rd),

K = span size,

The variance σ_C^2 of the new span (previous span plus or minus appropriate observations) takes the form

$$\sigma_C^2 = \frac{r}{K-3} \qquad (6)$$

where,

$$r = \Sigma Z_i^2 - [\Sigma Z_i \;\; \Sigma Z_i i \;\; \Sigma Z_i i^2] \; P \; \begin{bmatrix} \Sigma Z_i \\ \Sigma Z_i i \\ \Sigma Z_i i^2 \end{bmatrix}$$

$$K \geq 9$$
$$i = t - \left(\frac{K-3}{2}\right), \; t-\left(\frac{K-5}{2}\right), \; \ldots, \; t + \left(\frac{K+1}{2}\right)$$

Taking the ratio of σ_c^2 Eq. (6) and σ_v^2 Eq. (5), we obtain

$$R = \frac{20r}{\Sigma(\delta_i^3)^2} \;, \quad \begin{array}{l} K \geq 9 \\ i = t - \left(\frac{K-9}{2}\right), \; t - \left(\frac{K-11}{2}\right), \ldots, \; t + \left(\frac{K+1}{2}\right) \end{array} \tag{7}$$

If

$$R < F_{99} \; (K-3, \; K-3)$$

then we accept the corrected state estimate as being optimal and increase K by 2 and make the next prediction with Eq. (2).

However if

$$R > F_{99} \; (K-3, \; K-3)$$

then we can assume that probably the model (2nd degree polynomial in this case) no longer fits the input data, and decrease the span size by deleting the appropriate observation(s) from the span that has failed the F-test. In this case, we must make another correction with Eq. (3) before obtaining the next corrected state.

Obviously

$$X = \hat{X} + C$$

yields the smoothest state estimates. To obtain the filtered state estimates,
we have only to propagate the smoothed state estimates to the end point.

III. PERFORMANCE OF FIXED LEAST SQUARES AND DYNOSS

A. Noiseless Environment

1. Fixed Least Squares

The position estimates generated with fixed LS-51 are acceptable for the Athena trajectory. Figures 1 and 4 do not show any significant differences at this scale between the standard positions and those generated with LS-51.

On the Athena data, we note that the fixed LS-51 velocity estimates (Fig. 5) are close to the standard (Fig. 2) during intervals of time when there is little or no change in acceleration. However, this is not the case during abrupt changes in acceleration. In Fig. 5 note that during second stage ignition (approx. 46.0 sec), fixed LS-51 is in error by about 55 ft/sec. During second stage burnout (approx. 56.0 sec), the error is not as pronounced as before since the change in acceleration is more gradual than during ignition. The error in this case is around 30 ft/sec. The above observations can clearly be seen in the enlarged graphs (Figs. 6 and 7).

Examination of the Athena data shows that the fixed LS-51 acceleration estimates (Fig. 8) are acceptable during non-event periods. However during events, we note that the maximum error is approximately 120 ft/sec^2 (See the enlarged graphs, Fig. 9 and 10).

2. DYNOSS

The DYNOSS generated position estimates for the Athena are extremely close to the standard. Compare Fig. 4a to Fig. 1.

On the Athena velocity estimates, not only does DYNOSS show definite improvement over LS-51, but the maximum error is about 9 ft/sec. This error is noted only for one time point during ignition. Otherwise, the error is significantly less than 9 ft/sec. This can clearly be seen in Fig. 6. The excellent performance can be attributed to the decrease in the span size. Fig. 8 at bottom shows time vs. span size.

Comparison of DYNOSS generated acceleration estimates (Fig. 8) with the standard (Fig. 3), we notice close agreement at all times. We can also see from Figs. 9 and 10, the superior performance of DYNOSS during events.

B. Noisy Environment

1. Fixed Least Squares

The random noise introduced into the Athena position data was 50 feet, whereas the noise introduced into the High Acceleration data was 1 foot.

Although the Athena position estimates obtained from LS-51 do possess error, they are acceptable. Similarly the LS-61 position estimates on the High Acceleration data are acceptable. Compare Fig. 11 to Fig. 1 and Fig. 17 to Fig. 14.

With the introduction of 50 feet of noise to the Athena Standard positions, we notice in Fig. 12 that the velocity estimates generated by LS-51 have an oscillating pattern throughout the flight. The error is anywhere between 0 and 75 ft/sec. The nature of this kind of error usually cannot be tolerated, if a high velocity accuracy

is required for impact prediction.

The LS-61 velocity estimates on the 1 foot noise High Acceleration data (Fig. 18) demonstrates that during periods when the acceleration is not constant, there is a bias in the data. The magnitude and sign of the velocity bias depends on the rate of the acceleration change and its sign (compare LS-61 Fig. 18 with Fig. 15).

The undesirable pattern of the LS-51 acceleration estimates on the Athena for high noise levels, makes it unacceptable. The "random" acceleration estimates" generated by LS-51 are shown in Fig. 13.

Figure 19 shows the acceleration estimates propagated for the High Acceleration data corrupted with 1 foot random error. Note that the true minimums and maximums are never achieved with a fixed span size.

 2. DYNOSS

The position data generated by DYNOSS for both trajectories do possess slight errors but they are better than those from fixed least squares.

Comparison of DYNOSS to LS-51 (Fig. 12) velocity estimates indicates that DYNOSS is better. The effect noted in the LS-51 case has been avoided since DYNOSS has increased the span size to the maximum specified (See bottom Fig. 13). The main reason for the increase in span size can be attributed to the high noise content

of the data (50 ft). Also in Fig. 13, we note a slight decrease in span size due to second stage ignition and burnout.

Figure 18 reveals that the DYNOSS velocity estimates generated for the High Acceleration data with 1 foot of noise compare favorably with the standard (Fig. 15). Special note should be made that the bias obtained with LS-61 has been removed by using the optimal span size.
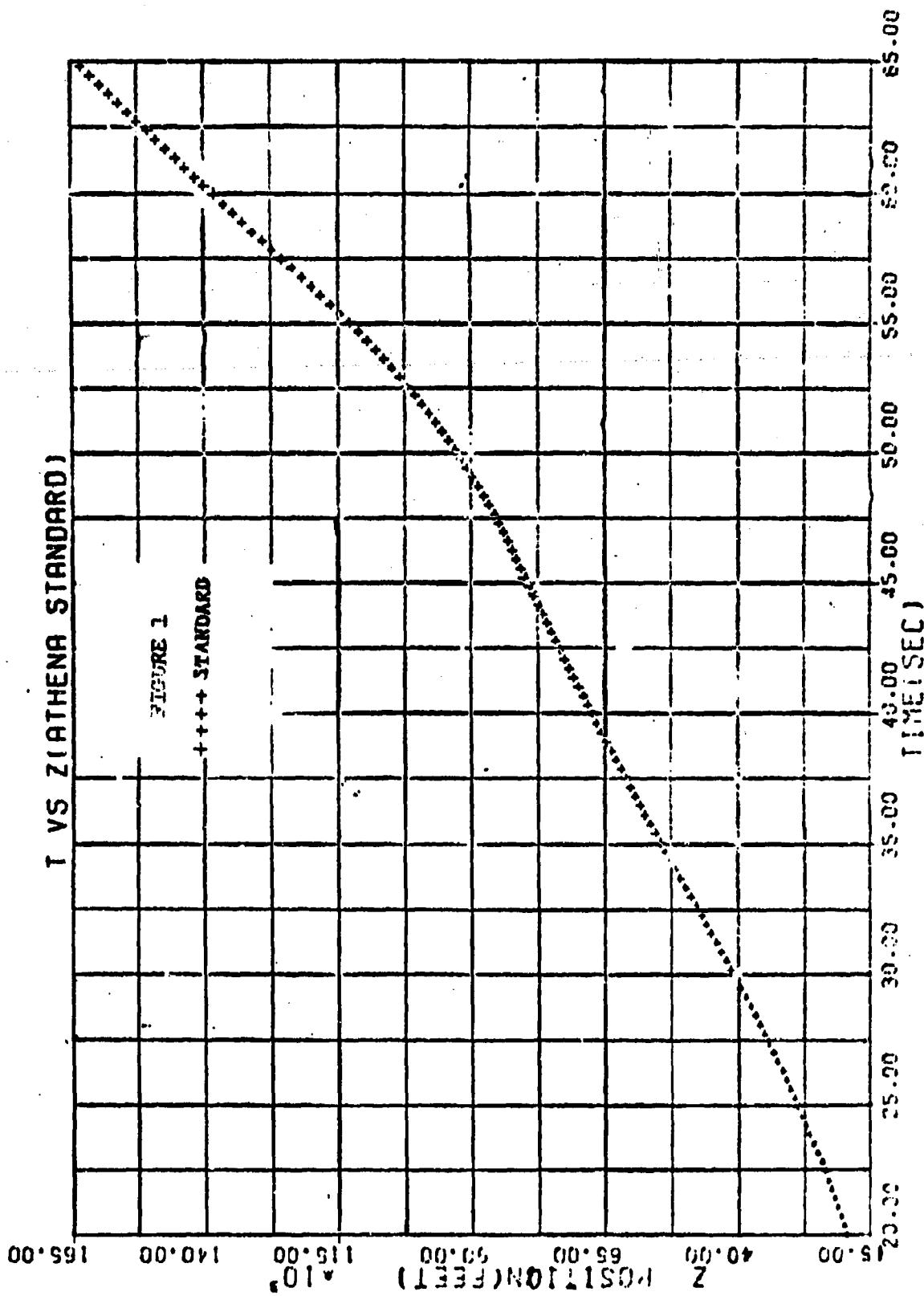
The use of DYNOSS to generate Athena acceleration estimates (Fig. 13) in a high noise environment has proven extremely helpful. Note that the harmful effect of using LS-51 on high noise level data to obtain acceleration estimates has been avoided. Comparison of DYNOSS to LS-51 (Fig. 13) shows the remarkable improvement in filter output when the span size is allowed to increase for non-event high noise level data.

The harmful effect of using too long a span (LS-61) on low noise (1 Ft) data to generate acceleration estimates can be seen in Fig. 19. Note the DYNOSS (Fig. 19) generated data is very close to the standard (Fig. 16).

C. Computing Efficiency

Figures 20 and 21 show a measure of the computer time required to generate DYNOSS and LS, respectively. The times indicated show the time required to input, compute, and output one smoothed point. In Fig. 20 maximum span means that the span was allowed (through load card options) to vary (depending on noise and events) between 9 points and the given maximum.

-233-

FIGURE 1

T VS Z(ATHENA STANDARD)

+++ STANDARD

Z POSITION(FEET) $\times 10^3$

TIME(SEC)
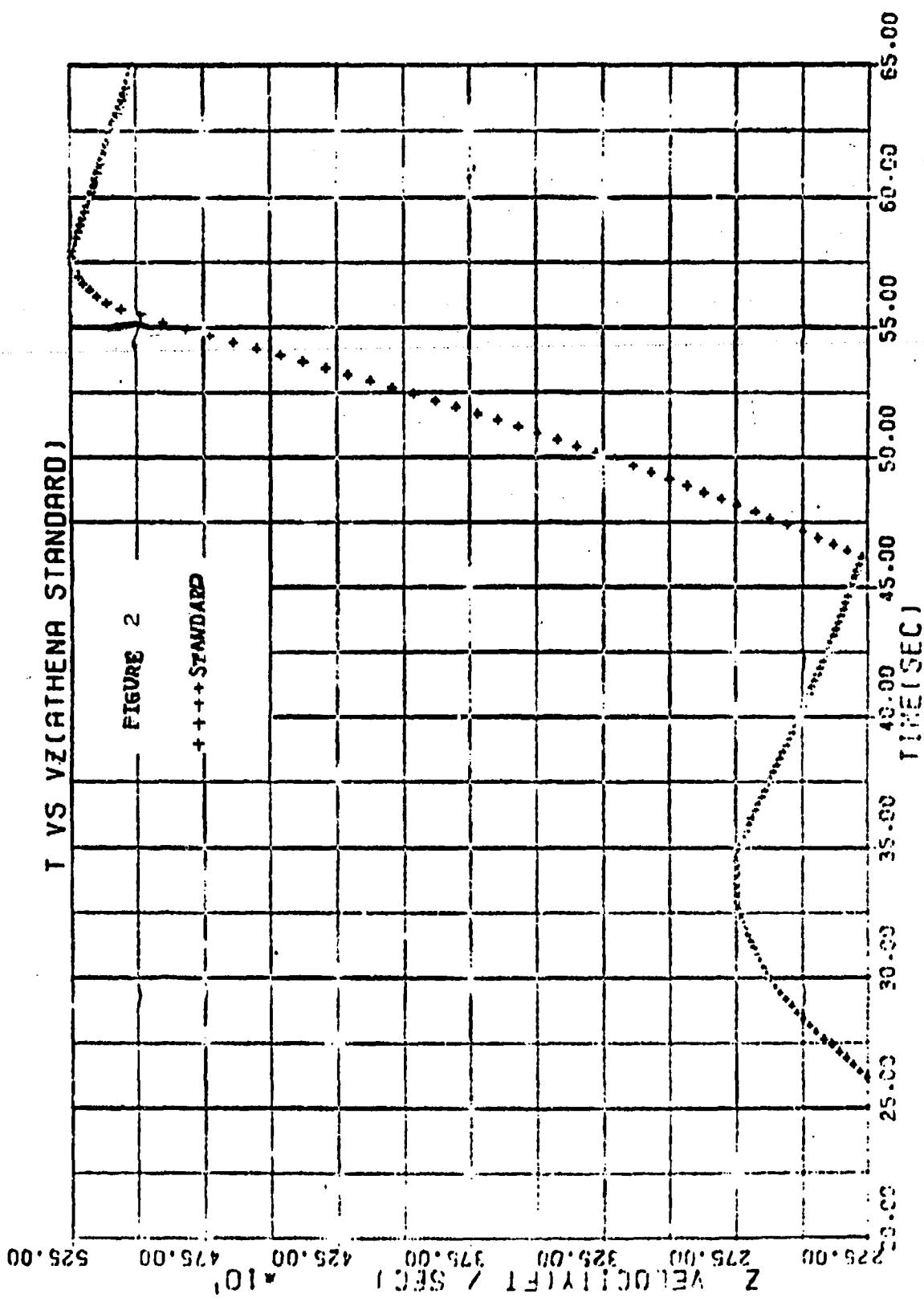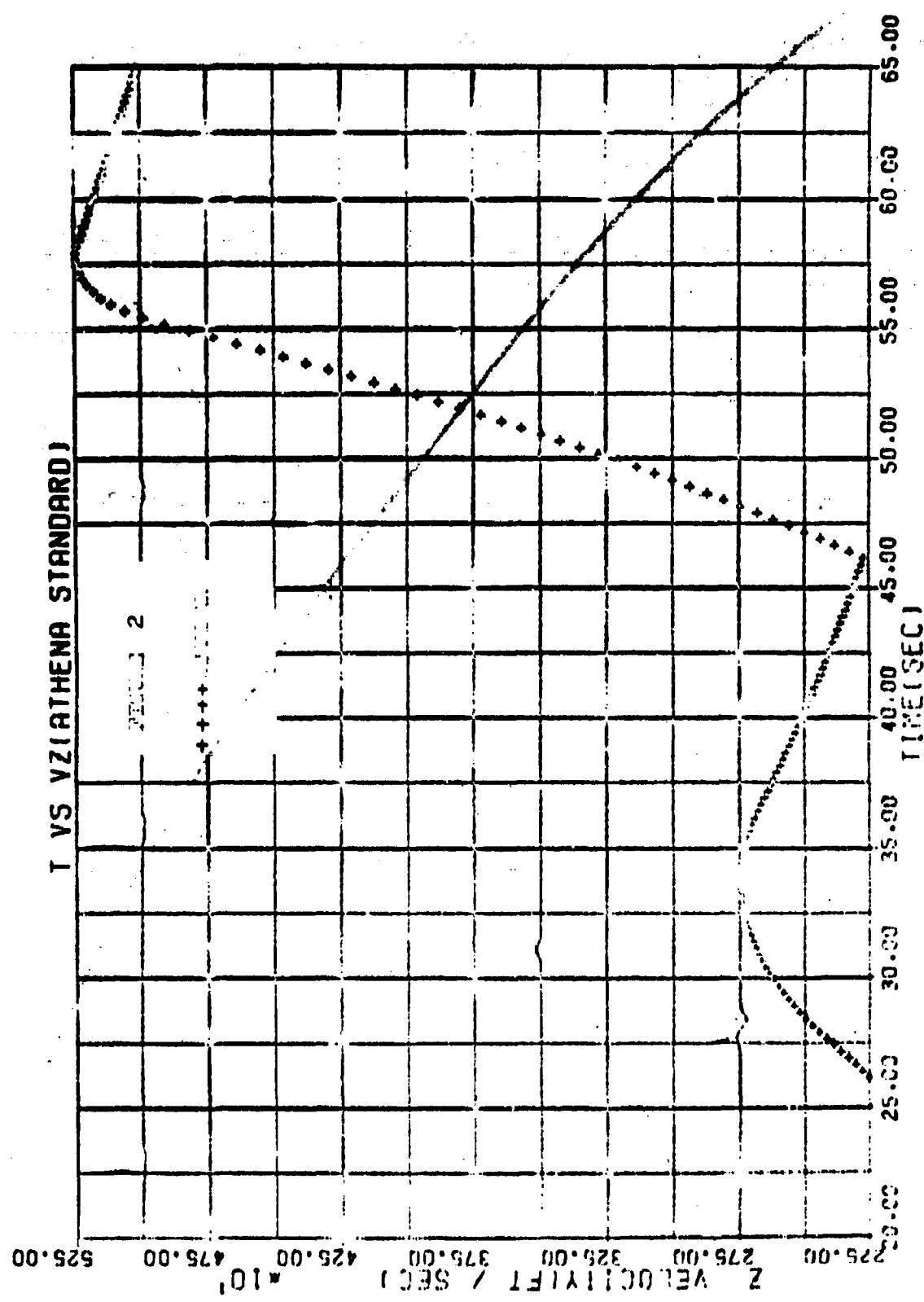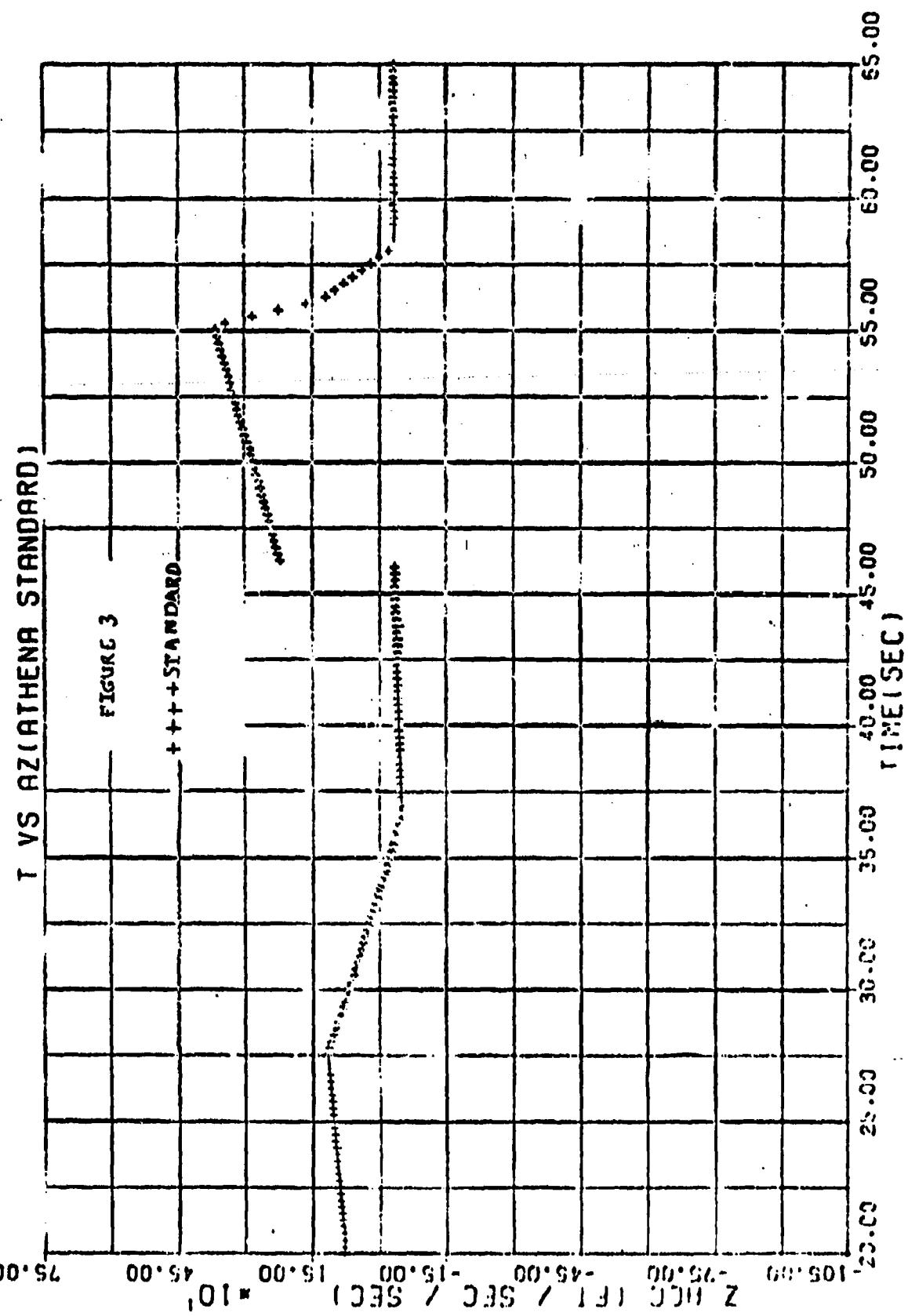
T VS VZ(ATHENA STANDARD)

FIGURE 2

+ + + STANDARD

Z VELOCITY(FT / SEC) *10¹

TIME(SEC)

-235-

T VS VZ(ATHENA STANDARD)

FIGURE 2

TIME(SEC)

Z VELOCITY(FT / SEC)

T VS AZ(ATHENA STANDARD)

FIGURE 3

+ + + STANDARD

TIME VS Z (NO NOISE)

FIG--- 4
..... IS-51

Z(FEET) #10³

TIME(SEC)

TIME VS Z (NO NOISE)

FIGURE 46

TIME VS VZ (NO NOISE)

FIGURE 5

DYNOSS ———

LS-51 ·····

VZ (FT/SEC) ×10¹

TIME(SEC)

TIME VS VZ (NO NOISE)

FIGURE 6

DYNOSS
LS-51
STANDARD

TIME VS VZ(NO NOISE)

FIGURE 7

—— DYNOSS
..... LS-51
+ + + STANDARD

TIME(SEC)

VZ(FT/SEC)

FIGURE 8

TIME VS AZ (NO NOISE)

DYNOSS
LS-51

FIGURE 9

TIME VS AZ (NO NOISE)

—— DYNOSS
.... LS-51
+ + + STANDARD

FIGURE 10

TIME VS AZ (NO NOISE)

PYNOSS
LS-51
+ + + STANDARD

-244-

TIME VS Z (50 FT NOISE)

TIME VS Z (50 FT NOISE)

FIGURE 12

TIME VS VZ (50 FT NOISE)

DYNOSS
········ LS-51

FIGURE 13

TIME VS AZ (50 FT NOISE)

FIG. 5-24

T VS VZ(HI ACC STANDARD)

FIGURE 25

+ + + + VZINTEG

V2 (FT/SEC)*10²

TIME(SEC)

T VS AZ(HI ACC STANDARD)

Figure 16

TIME(SEC)

AZ (FT/SEC/SEC) =10³

TIME VS Z (1 FT NOISE)

FIGURE 17

..... IL-61

-252-

TIME VS VZ (1 FT NOISE)

FIGURE 16

VZ (FT/SEC) ×10²

TIME(SEC)

-253-

TIME VS AZ (1 FT NOISE)

FIGURE 19

——— SIM
••••• IS-61

SPAN SIZE

-254-

DYNOSS COMPUTING TIME

FIGURE 20

* Using a UNIVAC 1108 Computer

LEAST SQ COMPUTING TIME

Figure 21

* Using a UNIVAC 1108 Computer

-256-

IV. **SUMMARY**

For noise free (except for round-off errors) data, DYNOSS was able to provide better state estimates than LS-51. This is readily apparent during events in the enlarged velocity and acceleration graphs (Fig 6, 7, 9, and 10).

With the introduction of 50 feet and 1 foot of noise to the Athena and High Acceleration data respectively, we noted DYNOSS performance was extremely good (Fig. 12 and 13). DYNOSS proved invaluable, especially for high acceleration non-event (not completely eventless, but for gradual thrusting periods) data when a long span creates an undesirable misfit in the LS-51 velocity estimates and acceleration estimates (See Fig. 18 and 19).

In our investigation, we also noticed that the DYNOSS span selector for noise free event data did in fact decrease the span size as expected during events (See bottom of Fig. 8) and therefore propagated optimal state estimates (Fig 6, 7, 9 and 10). Fig. 13 at the bottom shows the ability of DYNOSS to adjust to a noisy environment.

In Figs. 20 and 21 we noticed that DYNOSS is quite efficient and that the computer time is invariant with length of span.

We recognize that a considerably lower span size could have been used where noise content is low. Certainly this would have shown improvement in the LS velocity estimates (Fig. 18) from

1 to 2.33 sec. However, the estimates generated from 2.33 sec to 3.00 would then have deteriorated. Numerous other examples could be cited, but the important point is that fixed Least Squares _might_ propagate optimal state estimates only at certain times, whereas, DYNOSS always propagates optimal state estimates.

We hope that we have succeeded in giving adequate examples to point out the advantages afforded by DYNOSS.

V. CONCLUSIONS

One of the major conclusions reached is that dynamic filtering/smoothing is indispensable.

In a data reduction complex like Data Reduction Division at WSMR, New Mexico, DYNOSS is essential. The major reason is the diversity and enormous volume of data processed, frequently do not allow the data analyst to judiciously select the span size required to filter data. However, since acceleration and noise are non-stationary; not even the best analyst can select one span size to acquire optimal state estimates throughout a complete missile flight.

Thus, only by utilizing the optimal span size in any given time interval can a filter provide a true description of the missile dynamics.

In view of all the advantages afforded by DYNOSS and the test requirements for development of high accelerated multi-stage missiles, the implementation of dynamic filtering is not only necessary but mandatory.

The purpose of this appendix is to provide detailed computation of Eq. (4). Our first task is the computation of H. However H takes on different values depending on whether our span size is constant, increasing, or decreasing. When K is increased by two observations, we have

$$
H = \begin{bmatrix} 1 & \dfrac{(K-3)\Delta t}{2} & \left(\dfrac{(K-3)\Delta t}{2}\right)^2 \Big/ 2 \\[3mm] 1 & \dfrac{(K-1)\Delta t}{2} & \left(\dfrac{(K-1)\Delta t}{2}\right)^2 \Big/ 2 \end{bmatrix}.
$$

Decreasing K yields two values for H.

First if K is decreased by deleting two observations at the beginning of a span, we have

$$
H = \begin{bmatrix} 1 & \dfrac{-(K+1)\Delta t}{2} & \left(\dfrac{-(K+1)\Delta t}{2}\right)^2 \Big/ 2 \\[3mm] 1 & \dfrac{-(K+3)\Delta t}{2} & \left(\dfrac{-(K+3)\Delta t}{2}\right)^2 \Big/ 2 \end{bmatrix}.
$$

Second, if K is decreased by deleting two observations one at the beginning and the other at the end of a span yields

$$H = \begin{bmatrix} 1 & \frac{-(K+1)\Delta t}{2} & \left(\frac{-(K+1)\Delta t}{2}\right)^2 \Big/ 2 \\ 1 & \frac{(K+1)\Delta t}{2} & \left(\frac{(K+1)\Delta t}{2}\right)^2 \Big/ 2 \end{bmatrix}$$

The procedure to maintain a constant span makes use of two iterations of the basic WLSRE equations. During the first iteration, we increase the given span by adding an observation at the end. The second iteration will then decrease the span by deleting one observation at the beginning. In the former case, H is computed as follows

$$H = \begin{bmatrix} 1.0 & \frac{(K-1)\Delta t}{2} & \left(\frac{(K-1)\Delta t}{2}\right)^2 \Big/ 2 \end{bmatrix} .$$

In the latter case

$$H = \begin{bmatrix} 1.0 & \dfrac{-(K+1)\Delta t}{2} & \left(\dfrac{-(K+1)\Delta t}{2}\right)^2 \Big/ 2 \end{bmatrix}.$$

From the basic equation $\sigma^2 = (G + H^t WH)^{-1}$, σ_z^2 can be factored from G and W, leaving $(G + H^t WH)^{-1}$ as a function of K and Δt. Thus the $(G + H^t WH)^{-1}$ matrix does not have to be inverted every time but can be predicted based on the known values of K and Δt. It can be shown that

$$(G + H^t WH)^{-1} =$$

$$\sigma_z^2 \begin{bmatrix} \dfrac{3(3K^2-7)}{4K(K^2-4)} & 0 & \dfrac{-30}{K(K^2-4)\Delta t^2} \\ 0 & \dfrac{12}{K(K^2-1)\Delta t^2} & 0 \\ \dfrac{-30}{K(K^2-4)\Delta t^2} & 0 & \left(\dfrac{2}{\Delta t^2}\right)^2 \dfrac{12}{K(K^2-1)} \dfrac{15}{(K^2-4)} \end{bmatrix}$$

Once again we emphasize that K, is the K that takes into consideration the observation matrix Z.

Therefore, the correction takes the form

$$C = \sigma_Z^2 \begin{bmatrix} \dfrac{3(3K^2-7)}{4K(K^2-4)} & 0 & \dfrac{-30}{K(K^2-4)\Delta t^2} \\[3mm] 0 & \dfrac{12}{K(K^2-1)\Delta t^2} & 0 \\[3mm] \dfrac{-30}{K(K^2-4)\Delta t^2} & 0 & \left(\dfrac{2}{\Delta t^2}\right)^2 \dfrac{12}{K(K^2-1)} \dfrac{15}{(K^2-4)} \end{bmatrix} H^t W(Z-H\hat{X})$$

In the above equation if we substitute

$\dfrac{1}{\sigma_Z^2}$ for W we obtain

$$C = \begin{bmatrix} \dfrac{3(3K^2-7)}{4K(K^2-4)} & 0 & \dfrac{-30}{K(K^2-4)\Delta t^2} \\[3mm] 0 & \dfrac{12}{K(K^2-1)\Delta t^2} & 0 \\[3mm] \dfrac{-30}{K(K^2-4)\Delta t^2} & 0 & \left(\dfrac{2}{\Delta t^2}\right) \dfrac{12}{K(K^2-1)} \dfrac{15}{(K^2-4)} \end{bmatrix} H^t (Z-H\hat{X})$$

REFERENCES

1. Fierro, R., Nicholson C., "Comparison of Static and Dynamic Filters," Analysis and Computation Directorate, WSMR, New Mexico, October 1970.

2. Comstock D., Wright M., Tipton V., "Handbook of Data Reduction Methods," Data Reduction Division, WSMR, New Mexico, 1964, pp 355-368.

3. Comstock D., Introduction to Least Squares," Analysis and Computat 1 Directorate, WSMR, New Mexico, 1968.

AN "OPTIMIZER" FOR USE IN COMPUTER

SIMULATION: STULIES WITH A PROTOTYPE (U)

Dennis E. Smith
HRB-Singer, Inc.
Science Park, State College, Pennsylvania

ABSTRACT.

Because of their extreme complexity, many military operations research problems cannot be examined analytically, but instead must be examined by means of computer simulation. However, the use of simulation in attempting to locate an optimum solution often degenerates into a trial-and-error process. This paper describes a study aimed at providing the simulation user with efficient methods of searching for optimum computer simulation solutions.

This study, to date, has resulted in the following three major accomplishments:

(1) Development of a prototype "OPTIMIZER" computer program consisting of a set of search techniques.

(2) Empirically derived information to aid in the selection of a search technique.

(3) Successful application of the prototype "OPTIMIZER" to an existing computer simulation.

Each of these accomplishments is discussed in this report.

I. INTRODUCTION

Because of their extreme complexity, many military operations research problems cannot be examined analytically, but instead must be examined by means of computer simulation. When computer simulation is involved, an optimum solution is not obtained by analytic techniques, but rather by a search of the relevant parameter space. In general, this search for an optimum solution rapidly degenerates into a trial-and-error process. The goal of the present study is the development of more efficient procedures for obtaining optimum computer simulation solutions.

The introductory section of this report provides the necessary background information and summarizes the progress made toward the indicated goal. The remaining sections present the details of the various aspects of the current study effort.

A. BACKGROUND

A computer simulation may quite realistically be regarded as involving a "black box" in which the values of input variables or parameters are combined in some manner to produce output variables or responses. These responses are usually figures of merit of some type. Figure 1 summarizes this "black box" view of computer simulation when only one output variable (response) is involved.

As will be noted from this figure, the input variables are of two types: (1) controllable factors and (2) uncontrollable factors. The controllable factors are those input variables having values which can be controlled in the "real world." For example, input variables such as submarine speed, aircraft altitude, and destroyer bearing pertaining to United States Naval forces would be controllable factors because their values could be varied at will by the appropriate decision maker in the Naval chain of command. On the other hand, these input variables (submarine speed, aircraft altitude, and destroyer bearing) relating to enemy forces must be regarded as uncontrollable factors because the U.S. Navy cannot vary the values of these factors at will. In addition, environmental characteristics such as sea state, temperature, and amount of precipitation must also be regarded as uncontrollable factors because their values, too, cannot be controlled.

INPUT VARIABLES

CONTROLLABLE
FACTORS

$X_1,\ldots,X_k$

UNCONTROLLABLE
FACTORS

$Z_1,\ldots,Z_m$

MATHEMATICAL
MODEL
(COMPUTER PROGRAMS)

OUTPUT VARIABLE

RESPONSE

$y = g(X_1,\ldots,X_k | Z_1,\ldots,Z_m) + \epsilon$

(FIGURE OF MERIT)
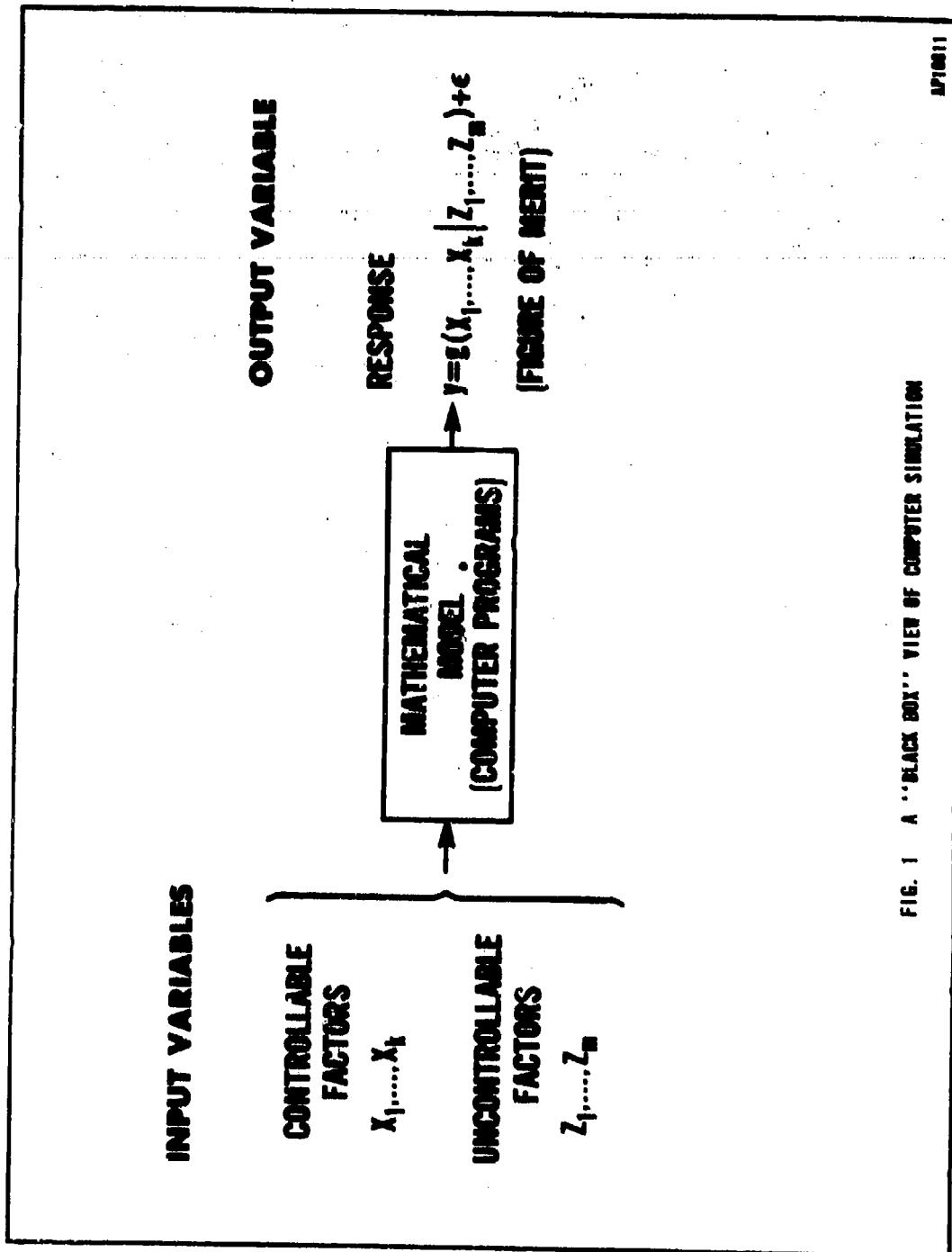
FIG. 1 A "BLACK BOX" VIEW OF COMPUTER SIMULATION

As Figure 1 indicates, the computer simulation "black box" produces the observed response y which is a function of both the controllable and uncontrollable factors. In probabilistic computer simulations where random variables are involved, the observed response y may vary even if two identical situations are run, i.e., if all input variables have the same values both times the simulation is used to produce a response.[1] In this situation, the observed response y may be assumed to vary about a true response $g(X_1, \ldots, X_k \mid Z_1, \ldots, Z_m)$, which, in statistical terms, is an expected value. The variation of y from this true response may be regarded as sampling error which arises from a random variable ϵ having zero expected value but positive variance.[2] Thus the observed response is composed of two quantities and may be expressed as $y = g(X_1, \ldots, X_k \mid Z_1, \ldots, Z_m) + \epsilon$, the form presented in Figure 1. It should be noted that the true response, g, defines a surface in k-dimensional space. This surface is known as a response surface.

The optimum solution consists of those values of the controllable factors which produce the true optimum response, i.e., the optimum expected value. It should be noted at this point that optimization, per se, in simulation is an ultimate goal which in practical situations is usually unreachable. Ordinarily, the most that one can hope to accomplish given the constraints of limited time and funds is the location of an improved, or a "good" solution. That is, the result may be the identification of a solution which, while not necessarily a global optimum, is acceptable and better than any previously available solution. To emphasize this fact in the current study, the term "optimization" when placed in quotation marks refers to the attempt to locate the true optimum solution. Thus, the "optimum" solution which is *found* by "optimization" may not be the true optimum solution.

B. SUMMARY OF FIRST YEAR'S EFFORT

HRB-Singer, Inc., under a contract[3] sponsored by the Office of Naval Research, has focused on "optimization" within that class of simulations where

---

[1] It is assumed, of course, that different sequences of pseudo-random numbers are used in each computer run.

[2] If the simulation is deterministic, however, ϵ is identically zero.

[3] Contract No. N00014-69-C-0285.

a single response (i.e., figure of merit) exists, the controllable factors are continuous (or at least approximately so) and there are no constraints on the controllable factors.

An HRB-Singer report[1], which was published in July 1970, describes the progress made during the first year of the ONR contract. This report, which discusses two categories of methods which may be used to aid in "optimization", concludes that the largest potential payoff would come from the development of an "OPTIMIZER" computer program. This "OPTIMIZER", which could be coupled to existing simulations, would automate, to a degree, the "optimization" process.

1. Methods to Aid in "Optimization"

Methods to aid in the quest for an "optimum" solution may be classified into two broad categories[2]: (1) internal methods and (2) external methods. Internal methods involve tinkering with the inner workings (the mathematical relationships and the computer programming) of the "black box". Because internal methods are actually built into the simulation when the mathematical models are constructed, the analysts and modelers responsible for model development must incorporate these methods directly into the "black box". External methods, on the other hand, do not affect the inner workings of the "black box", and are essentially independent of model construction. These methods specify search techniques which involve experimentation with different values of the controllable factors, usually using the output of the "black box" as feedback.

2. The Concept of an "OPTIMIZER" Computer Program

Smith (1970a) concluded that search techniques tend to provide the most useful means of "optimization", and that the existence of these techniques in an "OPTIMIZER" computer program should prove of value to the computer simulation user. Such an "OPTIMIZER" program could be constructed independently of any specific computer simulation, and would thus be general enough to provide flexibility and wide applicability. In addition, the user would not need an extensive

---

[1] Smith (1970a).

[2] Methods from each of these categories are discussed in Smith (1970a).

knowledge of the mathematics underlying the techniques since all computations could be handled within the structure of the "OPTIMIZER".

The "OPTIMIZER" would have to be designed in such a manner that it could be granted access to the inputs and outputs of the computer simulation, as well as to additional data[1] which the simulation user would supply. Figure 2 summarizes the design concept defining the ultimate goal of the current study, namely the eventual construction of a general "OPTIMIZER" program which can be tailored to any specific computer simulation situation by additional information supplied as input by the user.
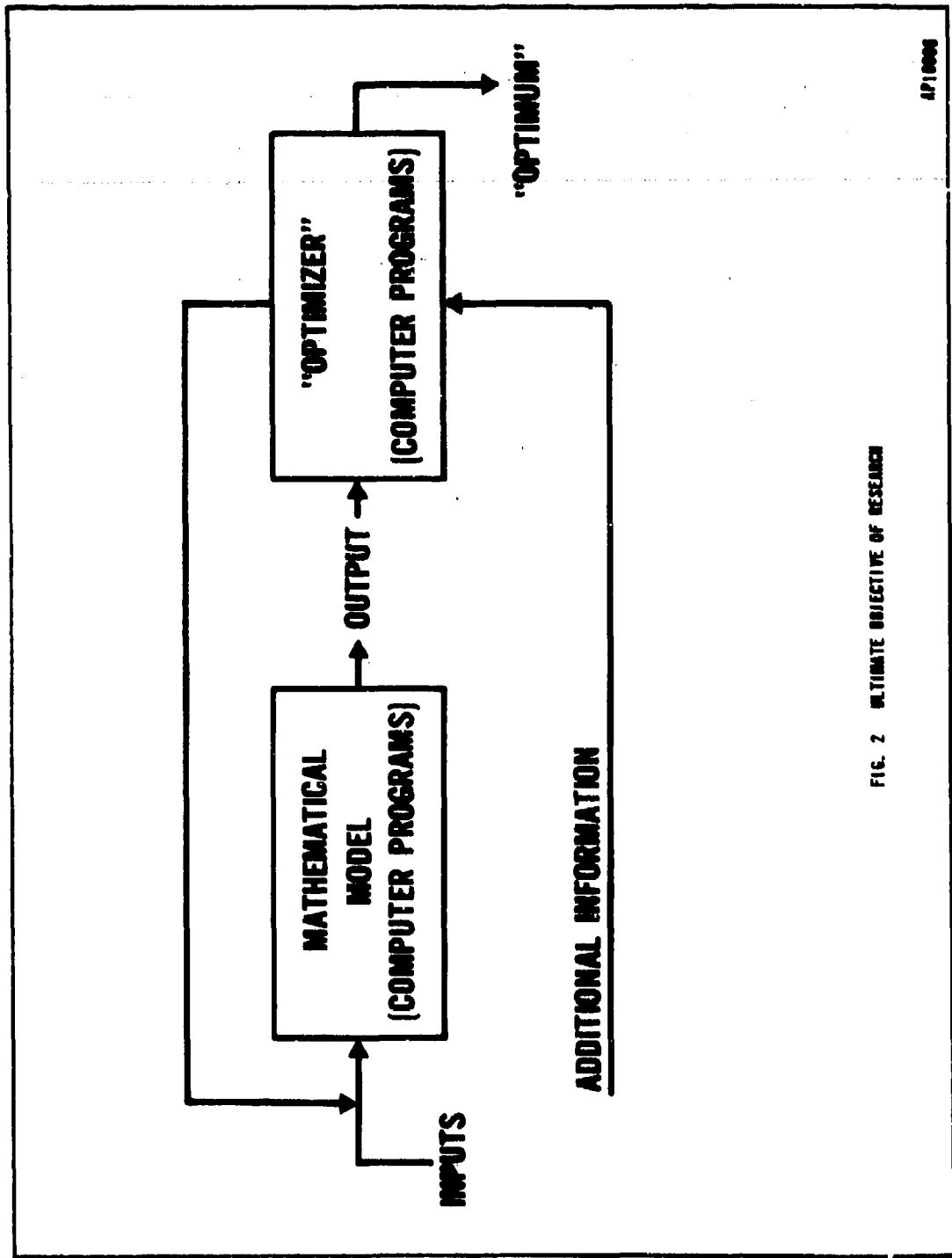
In essence, the "OPTIMIZER" would obviate the need for an expert who is familiar with the various search techniques and the underlying mathematics, because they would be contained within the "OPTIMIZER" itself. As indicated in Figure 2, the simulation user would be able to communicate to the "OPTIMIZER" information in addition to the standard simulation input. This information would be used by the "OPTIMIZER" to determine the best search technique to be adopted for the specific simulation situation. Thus, the "OPTIMIZER" as visualized would have a double optimization task. Not only would it attempt to find an "optimum" solution, but it would also attempt to determine the most appropriate "optimizing" technique to be used in any situation.

C. SUMMARY OF CURRENT STUDY EFFORT

During the past year, a prototype "OPTIMIZER" computer program has been programmed and checked for errors by running a number of test cases. This "OPTIMIZER" consists of seven specific search techniques. At the present time, selection of the particular technique to be applied in any given simulation situation must be made by the user of "OPTIMIZER". However, it is planned that the internal logic of the "OPTIMIZER" will eventually include decision criteria to permit automatic selection of the best technique for the existing simulation, based on user-supplied information about the simulation characteristics. To provide such decision criteria, an empirical study of the search techniques has been conducted using known response surfaces

---

[1] Such as the starting point for the search and the number of computer runs which are available.

FIG. 2 ULTIMATE OBJECTIVE OF RESEARCH

which were constructed to have various characteristics. In addition to these empirical studies, the prototype "OPTIMIZER" was applied to an existing computer simulation (i.e., to an unknown response surface).

Thus, the current year's effort consisted of three phases:

Phase I: Construction of a Prototype "OPTIMIZER"

Phase II: Empirical Examination of the Performance of the Search Techniques on Constructed Response Surfaces

Phase III: Application of the Prototype "OPTIMIZER" to an Existing Computer Simulation

Each of these three phases is discussed in the following sections.

1. The Prototype "OPTIMIZER"

The prototype "OPTIMIZER" was constructed to function as an executive program. As such, the "OPTIMIZER" provides linkages with the simulation, carries out the required bookkeeping, and performs the mathematical calculations connected with the search technique being used. In such a manner, the "OPTIMIZER" maintains control over the values of the controllable factors to be used at each step in the search for an "optimum".

The "OPTIMIZER" contains a total of seven search techniques, which are variations of three basic techniques. The first technique, Random Search, simply selects at random the values of the controllable factors to be used in each simulation run. The second technique, Single-Factor, examines one factor at a time by means of coordinated movement of a given factor while all others are held constant. The third technique, Response Surface Methodology (RSM), permits the examination of all factors simultaneously by means of an experimental design from which the estimated gradient direction may be determined.

2. Empirical Studies

As mentioned previously, the specific search technique to be applied must be selected by the user. In order for the user (or eventually, the "OPTIMIZER" logic) to choose the best technique, the performance of the

techniques must be evaluated. Because no mathematical basis exists for such an evaluation relative to characteristics which may vary from simulation to simulation, it is necessary to perform empirical studies in order to obtain the required information. A major portion of the current study was devoted to the construction of a framework for an empirical evaluation of the performance of the search techniques for different categories of simulation situations.

The evaluation itself was based on studies using response surfaces constructed from known mathematical relationships which were specifically chosen for the investigation. The mathematical relationships generating the response were so constructed that the true optimum was known, providing comparison of the performance of each technique with the true optimum, as well as comparison with the performance of other techniques. The performance of a technique, of course, is directly related to the "optimum" found by that technique. Specifically, the performance of a given search technique was measured in terms of the relative gain to be expected from using that technique, where relative gain was defined as:

$$\left\{ \frac{\begin{bmatrix} \text{True response at "optimum" point} \\ \text{found by search technique} \end{bmatrix} - \begin{bmatrix} \text{True response at} \\ \text{starting point} \end{bmatrix}}{\begin{bmatrix} \text{True optimum response} \end{bmatrix} - \begin{bmatrix} \text{True response at starting point} \end{bmatrix}} \right\}$$

Each of the search techniques was evaluated for a variety of simulation characteristics in the situation where the available number of computer runs was relatively small compared with the number of controllable factors. Specific characteristics considered were:

1. The number of controllable factors.

2. The number of computer runs available.

3. Distance from starting point to true optimum point

4. Relative size of the random error associated with the response surface.

5. Relative activity of the controllable factors.

6. The presence (or absence) of interaction.

7. The presence (or absence) of local optima.

Computer routines carried out the evaluation and provided summary tables and statistics of the performance of each of the techniques. As a result of this evaluation task, empirically derived payoff matrices for use in the selection of a search technique exist. These matrices, together with any existing knowledge of the simulation situation under consideration, permit the user to make a relatively efficient choice of a search technique.

3. Application of the "OPTIMIZER"

In addition to the research-oriented phase of the current study, HRB-Singer, with the cooperation of the U.S. Army Strategy and Tactics Analysis Group (STAG), applied the prototype "OPTIMIZER" to a computer simulation known as FORECAST II, which was developed by STAG. Although interface procedures were, in general, easily carried out[1], difficulties encountered with the controllable factors indicated that assumptions made for the prototype "OPTIMIZER" are not totally acceptable for many large military models.

In particular, the variables chosen as controllable factors in FORECAST II were neither continuous nor unconstrained, thus violating two assumptions on which the "OPTIMIZER" was developed. In an attempt to bypass the latter difficulty, a transformation of the controllable factors was used to alter the "optimization" problem into one which was unconstrained. Because the integer values which might be assumed by the controllable factors were not extremely small, it was decided to proceed by treating these factors as approximately continuous.

Results of the application proved positive. With only a small number of "OPTIMIZER" runs, definite improvement was achieved. This represented an increase in the observed figure of merit from .397 to .436. Because a theoretical upper bound to the figure of merit is .500, the relative gain provided by the "OPTIMIZER" may be calculated to be at least $\dfrac{.436 - .397}{.500 - .397}$

or approximately 38%.

---

[1] A fairly detailed account of the application of the "OPTIMIZER" to FORECAST II is contained in the Memorandum for Record of 25 February 1971 and 6 April 1971 prepared by Captain D. L. Renfro (formerly with Systems Development Division STAG).

II.   THE PROTOTYPE "OPTIMIZER"

The prototype "OPTIMIZER", which is comprised of a set of search techniques, was developed to function as a monitor or executive program. (As previously mentioned, the search technique to be used in any particular case must be specified by the user.) The linkage between "OPTIMIZER" and the simulation is via a FORTRAN CALL statement, with data exchanged via the CALL statement and COMMON statements.

The "OPTIMIZER" reads the initial (input) values for the controllable factors together with any other required input data, and then, in general, using observed response feedback, changes the values of the controllable factors according to the algorithms inherent in the prescribed search technique. In addition, the "OPTIMIZER" does the necessary accounting, maintaining a record of required information such as the number of runs executed and remaining, the results of calculations required by the search techniques, the best response observed, and the values of the controllable factors which have resulted in this best response.

The seven search techniques in the "OPTIMIZER" are based on three principal techniques[1]: (a) Random Search, (b) Single-Factor approach, and (c) Response Surface Methodology (RSM).   Of these three techniques, Random Search is a nonadaptive search technique, while the Single-Factor approach and RSM are adaptive search techniques.[2]

The prototype "OPTIMIZER" consists of the following seven specific search techniques:

[1] These basic search techniques are described in Smith (1970a). Another technique, the factorial design, was not included in the prototype "OPTIMIZER" because of the unwieldly number of computer runs it requires for even a small number of factors.

[2] In a nonadaptive technique, the observed response from a computer run does not influence the values of the controllable factors chosen for any succeeding run. An adaptive technique, however, bases selection of the values of controllable factors for future computer runs on the observed responses from previous runs.

(1) Random Search

(2) Single-Factor Approach

(3) Single-Factor Approach with Acceleration

(4) RSM - Variation I

(5) RSM - Variation I with Acceleration

(6) RSM - Variation II

(7) RSM - Variation II with Acceleration

The following sections provide descriptions of the basic search techniques[1] and the variations as they exist in the prototype "OPTIMIZER", and specify the input data required of the user. By far the most space is devoted to RSM because of its relative complexity. It will be assumed in the following descriptions that there are k controllable factors X_1, ..., X_k under investigation and that a maximum of $N = nm$ computer runs are available for use by the "OPTIMIZER", where m iterations are to be made at each point $(X_1, \ldots, X_k)$ in order to reduce statistical variation by obtaining average observed responses.

A. RANDOM SEARCH

An overall search region, defined by user input of upper and lower bounds on each of the k factors, is the region to which the search for the "optimum" is restricted. As its name implies, this search technique selects points at random from within the overall search region. Each of these k-dimensional points determines the values of the controllable factors to be used in obtaining an observed response. When all N computer runs have been used, the point which yielded the largest observed response is the "optimum" identified by Random Search.

If a user desires to use the Random Search option of the "OPTIMIZER", he must input the values of k, m, and n, and must also specify the upper and lower bounds for each controllable factor.

---

[1] Although these techniques are described in terms of a search for a maximum, they are equally applicable in a search for a minimum.

-276-

B. SINGLE-FACTOR APPROACH

The Single-Factor approach,[1] as its name implies, concentrates on one factor at a time. Specifically, when there are k controllable factors, $X_1, \ldots, X_k$, a starting point $(C_1, \ldots, C_k)$ is selected by the user, and a computer run made at this point. A step size Δ_1 is chosen and the X_1 value is changed to $C_1 - \Delta_1$ with all the other X_i's remaining at their starting values. A computer run is made at this new point. Similarly, a computer run is made at a point where $X_1 = C_1 + \Delta_1$ with the other X_i's remaining the same. If the observed response at either of these two points is better than at the starting point, the search is extended in the direction of the most improvement, and with all the other factors held constant, a new computer run is made corresponding to a value of $C_1 - 2\Delta_1$ or $C_1 + 2\Delta_1$ for X_1, as is appropriate. The search continues in the same direction as long as improvement continues to be made.

If a point is reached at which improvement is no longer being made, the best previous point is regarded as a new starting point, a step size Δ_2 is chosen, and the process is repeated, changing only X_2 while keeping the values of the other factors fixed. After factor X_k is examined, the cycle is completed by then considering X_1 again. At the end of a cycle, each Δ_i is halved before a new cycle begins. This procedure continues with each factor, in turn, until either the process is terminated by the identification of an "optimum" point, or the number of computer runs is exhausted.

In addition to the values of k, n, and m, the user of the Single-Factor option of the "OPTIMIZER" must input the initial values of the controllable factors, i.e., the starting point $(C_1, \ldots, C_k)$, and an initial step size Δ_i for each factor X_i, $i = 1, \ldots, k$.

C. RSM

Unlike the Single-Factor approach which varies only one factor at a time, RSM[2] permits the controllable factors to be varied simultaneously. Thus, although the direction of search movement always remains parallel to the X_i axes in a Single-Factor search, the RSM search allows movement in any direction

---

[1] Developed by Friedman and Savage (1947).

[2] Developed by Box and Wilson (1951)

However, data input for a RSM option is the same as that for the Single-Factor option. That is, the user must input the values of k, n, and m, the starting point $(C_1, \ldots, C_k)$ and a step size Δ_i for each factor X_i, $i = 1, \ldots, k$.

RSM uses an experimental design in a local region about the starting point to determine the estimated gradient direction known as a path of steepest ascent. Computer runs corresponding to points on this path, which provides the direction of predicted optimum response, are made until there is no improvement in the observed response, at which time the whole process is repeated within a smaller experimental region. At a later stage in RSM, additional experiments based on second order designs[1] may, if desired, be conducted to determine the approximate curvature of the response surface in the vicinity of a given point. However, because of the large number of computer runs required, this step will rarely be used in dealing with practical simulation problems. Thus, only the first phase of RSM (the initial experiment and the exploration of the path of steepest ascent/ have been included in the prototype "OPTIMIZER".

In general, the calculations required by RSM are simpler if made in terms of transformed, or coded, factors $x_1, \ldots, x_k$, where $x_i = (X_i - C_i)/\Delta_i$ for $i = 1, \ldots, k$. In terms of the coded factors, the starting point of the search is at $(0, \ldots, 0)$ and the initial step size is unity for each of these k factors. In the following discussion it will often be more useful to refer to the coded factors instead of the original factors.

1.   The Initial Experiment

The two RSM variations which are included in the prototype "OPTIMIZER" differ only in the initial experiment that is used to determine the path of steepest ascent. Variation I of RSM uses a 2^k factorial or a 2^{k-p} fractional factorial[2] of minimal size as the initial experiment, while Variation II uses a simplex design[3] involving $k + 1$ computer runs, the minimum required for calculation of a path of steepest ascent.

---

[1] For a discussion of second order designs, see Cochran and Cox (1957).

[2] See Cochran and Cox (1957) or Davies (1967).

[3] See Box (1952).

Only two values of each factor X_i are used in the initial experiment of Variation I. These values define a local region about the starting point $(C_1, \ldots, C_k)$ and consist of an upper value $U_i = C_i + \Delta_i$ and a lower value $L_i = C_i - \Delta_i$. The experiment is constructed under symmetry constraints which guarantee that each of the upper and lower values of X_i will be used an equal number of times. In addition, each of these two values of X_i appears the same number of times with the upper and lower values of each other factor. For example, the computer runs dictated by the initial experiment for the case $k = 3$ would be:

$$(L_1, \; L_2, \; U_3) = (C_1 - \Delta_1, \; C_2 - \Delta_2, \; C_3 + \Delta_3)$$

$$(U_1, \; L_2, \; L_3) = (C_1 + \Delta_1, \; C_2 - \Delta_2, \; C_3 - \Delta_3)$$

$$(L_1, \; U_2, \; L_3) = (C_1 - \Delta_1, \; C_2 + \Delta_2, \; C_3 - \Delta_3)$$

$$(U_1, \; U_2, \; U_3) = (C_1 + \Delta_1, \; C_2 + \Delta_2, \; C_3 + \Delta_3)$$

In terms of the coded factors, the initial experiment would be:

$$(-1, \; -1, \; +1)$$

$$(+1, \; -1, \; -1)$$

$$(-1, \; +1, \; -1)$$

$$(+1, \; +1, \; +1)$$

It can easily be seen in this case that the symmetry conditions hold.

This symmetry, inherent in the experiments used in Variation I, provides efficient estimation of gradient direction by permitting the effect of a given factor X_i to be estimated by averaging over a number of combinations of the other factors. Thus, these experiments provide for internal replication which, in general, allows for accurate estimation by reducing the size of the statistical error associated with the required estimates. The major drawback to the experiments used in Variation I is that they require a number of computer runs which is a power of two. Because a minimum of $k + 1$ computer runs is required to estimate the path of *steepest* ascent when k factors are involved,

one would, in general, not wish to spend many additional runs in initial experimentation. If $k = 35$, for example, the initial experiment in RSM - Variation I would require 64 computer runs, or 28 more than the minimum required.[1]

Unlike Variation I, Variation II of RSM is based on an initial experimen which always uses the required minimum number of runs to obtain the path of steepest ascent. This initial experiment results from the simplex design discussed by Box (1952). The nomenclature used in describing the design arises from the fact that, viewed geometrically, the design is formed by the vertices of a k-dimensional simplex.

Although the simplex design can be constructed to provide an estimate of gradient direction which is as accurate as that provided by the experiment of Variation I, this construction is extremely dependent upon the assumption that the response surface can be approximated to a good degree by a hyperplane, i.e., a linear function, in a local region about the starting point. Thus, implicit in Variation I is the assumption of an approximating hyperplane within the locality around the starting point defined, in terms of the coded factors, by $x_i = \pm 1$ for $i = 1, \ldots, k$. A similar assumption is implicit in Variation II. However, for Variation II to provide an estimate of gradient direction with the same accuracy as the initial experiment of Variation I, it turns out that computer runs must be taken corresponding to some values of the x_i's which are quite distant from the interval $(-1, +1)$. For example, at least one of the x_i's must assume the value of $\sqrt{k}$. Thus, for $k = 100$, a computer run must be taken for $x_i = 10$. This unhappy circumstance arises from the orientation of the points of the simplex design and is a result of the "vastness" of hyperspace. which was discussed by Hooke and Jeeves (1958) In actuality, each point of the simplex design and each point of the fractional factorial are exactly the same distance from the starting point in k-dimensional hyperspace. However, a point in the fractional factorial with coordinates which assume the values of -1 and +1 is nearer the starting point in an individual coordinate sense than is a point in the simplex with coordinates which may assume the values of $-\frac{1}{\sqrt{k}}$ and $\sqrt{k}$. Although none of these coordinates is too far distant from the interval

---

[1] Actually, the situation is not quite as bad as it might seem. This point is discussed further in Section III. D.

-280

(-1, +1) when k is relatively small, the distance becomes intolerable when k is large. Therefore, the simplex design in the prototype "OPTIMIZER" was modified to insure that any value which a given coded factor x_i must assume is not farther than $\sqrt{2}$ from the starting value 0. Such modification, however, tends to result in a less accurate estimate of gradient direction.

In general, then, Variation II provides an initial experiment containing less runs than the initial experiment of Variation I, while Variation I estimates gradient direction more accurately than does Variation II. However, once the path of steepest ascent is determined, the same procedure is followed by both of the variations.

2. Following the Path of Steepest Ascent

In addition to the computer runs specified by the initial experiment, a run is also made at the starting point. Based on the observed responses obtained from these computer runs, the "OPTIMIZER" determines the path of steepest ascent. Using the largest observed response, the "OPTIMIZER" calculates the point on the path which provides a predicted response with the same value, and makes a computer run at this point. (This is the first computer run taken on the calculated path.) Another computer run is then automatically taken on the path[1] a distance Δ away from the first point (in the coded factors). Because preliminary studies indicated that setting $\Delta = .5\sqrt{k}$ provides fairly good results,[2] this is the value of Δ in the "OPTIMIZER". However, if the user should decide that another value for Δ would be more appropriate, he has the option of inputting that value.

After the two initial computer runs are taken on the path of steepest ascent, additional runs are taken on the path, each one a distance Δ away from the previous run, so long as an improved response is observed. If a point is

---

[1] This second computer run on the path is taken regardless of whether or not the first computer run provided improvement. This procedure guards against the possibility that the first point was not far enough out on the path.

[2] It should be noted that in a previous study, McArthur (1961) found that $\Delta = 1.2$ appeared to give the best results for cases involving values of k ranging from 2 to 10. This value of Δ falls into the interval given by $.5\sqrt{k}$ for k = 2 and k = 10, i.e., the interval (.707, 1.5), lending some credence to the present choice of the value of Δ.

reached at which the observed response is less than a previous observed response, the "OPTIMIZER" uses the point corresponding to the maximum observed response as the center of a new experiment. In terms of the original factors $X_1, \ldots, X_k$, the center point has moved from $(C_1, \ldots, C_k)$ to a new point. When this occurs, a new step size for each factor X_i is defined as one-half the existing step size Δ_i.

At such a juncture, the "OPTIMIZER" determines whether or not a sufficient number of runs remain to allow all the runs dictated by another experiment to be carried out. If so, the RSM search proceeds in a manner similar to that previously described. If not, the "OPTIMIZER" determines the largest number of factors which can be included in a new experiment within the remaining number of runs and still leave a minimum of two runs on the path of steepest ascent. The "OPTIMIZER" uses this reduced set of k^* ($<k$) factors to calculate a new direction of steepest ascent. The reduced set of factors consists of those k^* factors corresponding to the k^* factors found by the previous experiment to have the largest effect on the path of steepest ascent.

This procedure is repeated until either the specified number of computer runs are exhausted or the center point of an experiment provides a better response than any of the other points in the experiment. When either of these events occurs, the best observed point is labeled the "optimum". The latter event indicates convexity of the surface in the vicinity of the center point, implying that application of a second-order design[1] would be warranted. The prototype "OPTIMIZER", however, was developed under the assumption that this stage in the search process would rarely, if ever, be reached, which follows implicitly from the assumption of an extremely limited number of available computer runs.

D.   THE ACCELERATED TECHNIQUES

The accelerated versions of Single-Factor, RSM-Variation I, and RSM-Variation II operate in a manner similar to the unaccelerated versions. However, when movement in a given direction (e.g., in the positive X_i direction

---

[1] See, for example, Davies (1967) or Cochran and Cox (1957).

or along a path of steepest ascent) provides improvement in the observed response at three successive points, the accelerated versions use an algorithm to speed up movement in that direction. To simplify the ensuing discussion, only RSM acceleration on the path of steepest ascent will be described. The procedure for Single-Factor acceleration is similar.

If the assumption of an approximating hyperplane is reasonable, it can be shown[1] that an estimate $\hat{y}$ of the value of the response at the coded point $(x_1, \ldots, x_k)$ is given by

$$\hat{y} = b_o + \sum_1^k b_i x_i$$

where the b_i's are estimates obtained from the experiment. The estimates $(b_1, \ldots, b_k)$ determine the direction of the path of steepest ascent in terms of the coded factors and thus, any point on the path may be denoted by $(wb_1, \ldots, wb_k)$. The predicted response at this point is given by

$$\hat{y} = b_o + \sum_1^k b_i (wb_i)$$

$$= b_o + w \sum_1^k b_i^2 .$$

Now, as described in the previous section, each point on the path is a distance of Δ from the previous point. Therefore, if the initial point on the path corresponds to $w = w_1$, the second point on the path then corresponds to $w = w_2 = w_1 + (\Delta / \sqrt{\sum_1^k b_i^2})$. In general, the jth point on the path corresponds to $w = w_j = w_1 + (j - 1) (\Delta / \sqrt{\sum_1^k b_i^2})$.

---

[1] See, for example, Davies (1967) or Cochran and Cox (1957).

Thus, the distance between the point corresponding to w_j and the point corresponding to w_{j-1} is

$$\sqrt{\sum_1^k (w_j b_i - w_{j-1} b_i)^2} = \sqrt{\sum_1^k b_i^2 \, (w_j - w_{j-1})^2}$$

$$= \sqrt{\sum_1^k b_i^2} \; \sqrt{\Delta^2 / \sum_1^k b_i^2}$$

$$= \Delta ,$$

as required. By means of some algebraic manipulation, the predicted response at the jth point on the path may be written as

$$\hat{y}_j = j A_0 + B_0$$

where

$$A_0 = \Delta \sqrt{\sum_1^k b_i^2}$$

and

$$B_0 = b_0 + (w_1 - \Delta / \sqrt{\sum_1^k b_i^2}) \sum_1^k b_i^2$$

Thus, the predicted response is a linear function of j.

Again, this result is contingent on the assumption of an approximating hyperplane. However reasonable this assumption within the starting locality, it quickly loses credibility as points further out on the path are considered. A bit more general, and hence more plausible, assumption might be that the predicted response is approximately a quadratic function of j. That is, the assumption might be made that

$$\hat{y}_j = j^2 A_1 + j B_1 + C_1$$

Based on this assumption, the observed responses from the first three points on the path of steepest ascent may be used to estimate the values of A_1, B_1, and C_1. In addition, the estimated location of an extremum can be obtained from $j = \dfrac{-B_1}{2A_1}$, where

$$A_1 = .5 (y_3 - 2y_2 + y_1)$$

and

$$B_1 = 4y_2 - 1.5y_3 - 2.5y_1$$

An acceleration algorithm, based on the preceding discussion, was developed for optional use within the "OPTIMIZER". Specifically, this algorithm contains the following five steps:

(1) Observe the response y_1 at the initial point on the path and the response y_2 at the second point on the path.

(2) If $y_2 > y_1$, observe response y_3 at the third point on the path.

(3) If $y_3 > y_2 > y_1$, compute the coefficients A_1 and B_1 and determine $j = \left[\dfrac{-B_1}{2A_1} \right]$, that is, the truncated value of $\dfrac{-B_1}{2A_1}$.

(4) If $j > 11$ or $j \leq 1$, set $j = 11$. Otherwise, retain the calculated value of j. (A value of $j \leq 1$ provides as an extremum the minimum of a quadratic function which is concave upward, reflecting the fact that the observed responses are increasing at an increasing rate. In this situation, setting $j = 11$ yields a point which is a reasonable distance on the path in view of the fact that the quadratic assumption may not hold and the observed responses may contain random error. Similarly, if $j > 11$, setting $j = 11$ guards against venturing too far out on the path.)

(5) Observe the response y_j corresponding to the jth point. If $y_j > y_3$, repeat the preceding steps using y_j as the starting point. If $y_j \leq y_3$, backtrack on the path to get y_{j-1}. If $y_{j-1} \leq y_j$, return to y_3 and continue on the path in the standard manner. If, however, $y_{j-1} > y_j$, observe y_{j-2}, etc., as long as improvement continues, and when improvement stops, say at y_h, and if $y_h > y_3$, perform a new experiment with the hth point as center. If $y_h \leq y_3$, however, go back to y_3 and continue on the path in the standard manner.

Similar steps are followed for the accelerated version of the Single-Factor approach, the only difference being that only one factor is changed at a time and the steps are made in terms of a Δ_i change for factor X_i, which corresponds to a change of one unit in the coded factor x_i.

III. EMPIRICAL INVESTIGATION OF THE SEARCH TECHNIQUES

Because no analytical basis exists for evaluation of the performance of the search techniques relative to characteristics which may vary from simulation to simulation, it is necessary to perform empirical studies in order to obtain a comparison of the techniques. Although such empirical studies could be based on unknown response surfaces generated by actual computer simulations, the current research effort used response surfaces which were constructed from known mathematical relationships specifically chosen for the investigation. This procedure has many advantages associated with it. Response surfaces constructed in this manner present difficulties of the same *type* as difficulties encountered in a full-scale simulation. Because full-scale simulations are not required in the evaluation, however, there is a substantial saving in costs (associated with modeling, programming, and computing). Also, the known mathematical relationships generating the response surface can be so constructed that the true optimum is known, thus permitting a comparison of the "optimum" found by a search technique with the actual optimum, in addition to a comparison with the "optimum" found by any other technique.

Although empirical investigations of search techniques on known response surfaces were carried out by Brooks (1959) and McArthur (1961), they considered cases where only a few controllable factors were involved. (The majority of cases considered were two-factor problems.) Although Brooks and McArthur used different criteria for evaluating the performance of the search techniques, their conclusions agreed. These conclusions were, in general, to use RSM if only a small number of factors were involved, and to use Random Search if a large number of factors were involved. Although Brooks did not give any definite value for distinguishing between a small number and a large number of factors, McArthur's study indicated that in any case involving more than three factors, Random Search should be used.

The current study has benefited from both of these previous investigations, particularly from McArthur's work, which used ten properties to characterize optimization search problems. A similar classification method, described in the following sections, was used in the current effort.

A. CRITERIA USED IN THE INVESTIGATION

As a first step in planning for an investigation of the performance of the various search techniques, the properties suggested by McArthur were revised to serve as a basis for categorizing simulation situations. This revision resulted in the following list of characteristics:

1. The number of controllable factors.

2. The number of available computer runs.

3. The presence or absence of local optima.

4. The size of the random error (statistical variation) present.

5. The distance of the starting point of the search from the true optimum point.

6. The relative activity of the factors.

7. The presence or absence of interaction among the factors.

The evaluation of search technique performance involved the application of each of the techniques on response surfaces constructed to have specific combinations of these characteristics.

In order to normalize the measure of performance of the techniques on the various resultant surfaces, it was decided to use the concept of relative gain based on true responses rather than on observed responses. (A true response is the response which would have been observed if no random error were present.) Specifically, if

(a) R_F denotes the true response at the "optimum" found by a search technique,

(b) R_S denotes the true response at the starting point of the search, and

(c) R_O denotes the true optimum response,

then gain will be defined as $(R_F - R_S)/(R_O - R_S)$. That is, the gain is a fraction of the maximum possible improvement, $R_O - R_S$, that could be made. It should be noted that although the gain can never be greater than 100 percent, it is

conceivable that it may be negative in some cases. A negative gain might result, for instance, because of the presence of random error which resulted in a search technique identifying an "optimum" point which actually yields a true response less than that of the starting point.

In light of this definition of gain and the various simulation characteristics mentioned previously, the evaluation of search technique performance attempted to provide at least a partial answer to one main question. Specifically, this question is: if k controllable factors are under consideration and N computer runs are available, what is the best search technique to use and how much gain is to be expected from using this technique on response surfaces having various characteristics?

The specific aspects of the characteristics chosen for inclusion in this study are delineated in the following sections.

1. The Number of Controllable Factors

The number of input variables for computer simulations of military operations research problems tends to be extremely large. Although only a subset of these variables will be included as controllable factors in any particular optimization problem, the number of factors in this subset can easily be 20, 120, or even larger. Very rarely does one see, in the simulation framework, an optimization problem involving only a few (e.g., less than ten) controllable factors.

To permit an examination of the performance of the search techniques in situations involving a reasonable number of controllable factors, it was decided to deal with 30-factor and 120-factor problems. These values provide at least rough bounds for the number of factors a simulation user is willing to consider in a trial-and-error or heuristic search for an optimum solution. Thus, in the current study, the values of k (the number of controllable factors) considered were k = 30 and k = 120.

This choice of values of k permits the use of fractional factorials which require a number of computer runs which is only slightly larger than the value of k. For example, a value of k = 120 permits the use of a fractional factorial involving 128 computer runs, while a value of k = 130, for instance, would require the use of a fractional factorial involving 256 computer runs. This

increase of 128 required computer runs corresponding to an increase of only 10
in the value of k results because of the manner in which the fractional factorials
are constructed[1]. Implications of the choice of the values of k for inclusion in
this study are discussed in Section III.D.

2. The Number of Available Computer Runs

In general, the simulation user knows, at least roughly, how much com-
puter time he may expend in a search for an optimum solution. Of course, the
user usually regards this amount of time as insufficient. Nonetheless, it will
implicitly be assumed that the user will do the best he can within the allotted
time. Because the running times of various simulations are inherently different,
to say nothing of the differences in the speed of different digital computers, it is
more convenient to deal with the number of available computer runs, N.

In the typical case, the available number of computer runs is relatively
small compared with the number of controllable factors. In view of this, the
cases considered in the present study involved $N \leq 2k$ computer runs, and im-
plicitly assumed that only one iteration would be made at a given point. Specifi-
cally, the performance of each search technique was evaluated for values of
$N = .5k$, $1.1k$, $1.5k$, and $2.0k$. In other words, if a 30-factor problem were
being considered, a given search technique would locate an "optimum," first
with a maximum of 15 runs, next with a maximum of 33 runs, next with a maxi-
mum of 45 runs, and finally, with a maximum of 60 runs.

3. Presence or Absence of Local Optima

The response surfaces generated by computer simulations may be of
many different varieties. In general, a simulation user searching for an optimum
hopes he is facing a situation in which there is a global optimum but no local
optima. In other words, he is hoping for a unimodal response surface, because
if he does find an optimum, he knows that it is the global optimum. A much
worse case is one in which numerous local optima exist in addition to the global
optimum. In this situation, the user is facing a response surface which consists
of various peaks, valleys and saddlepoints. Hence, if he does locate an optimum,

---

[1] See the discussion of the situation in Section II.C.

he has no assurance that it is not merely a local optimum which provides a response much worse than that of the global optimum. Figure 3 illustrates a unimodal response surface consisting of only a global optimum and no local optima, while Figure 4 illustrates a multimodal response surface consisting of a global optimum and many local optima. Although both of these surfaces are defined by only two factors to permit illustration, the extension to a larger number of factors is straightforward, but essentially impossible to visualize. Because there are many ways in which to construct unimodal and multimodal surfaces, the specific mathematical forms giving rise to the surfaces used in the study will be briefly described here. The actual construction of these surfaces is discussed in Section III.B.

The unimodal surface which possesses only a global optimum was generated by a quadratic equation in the k factors. The equation giving rise to this type of response surface was, in generic form,

$$y = -\sum_{1}^{k} A_i \left[Z_i (X_1, \ldots, X_k) \right]^2 \quad -$$

where the Z_i's are linear functions of the X_i's. The specific form of the functions Z_i is dependent on the presence or absence of interaction which is discussed in Section III.A.7, and the A_i's assume the values 0, 1, or 4 as determined by the relative activity of the factors, as described in Section III.A.6.

The multimodal surface having local optima was generated by a combination of cosine terms damped by exponential terms. The equation used in generating the response surfaces was, in generic form,
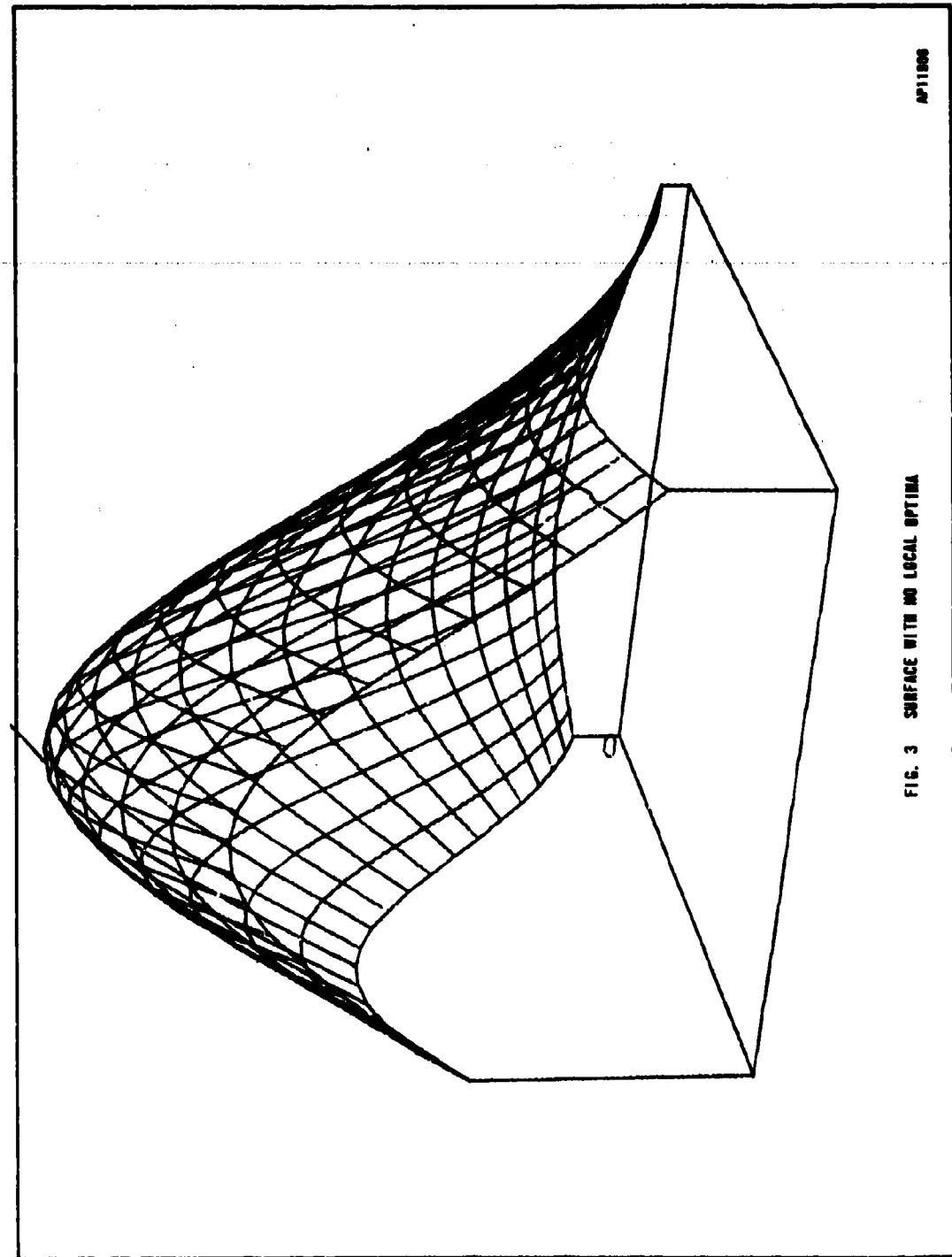
$$y = -\sum_{1}^{k} A_i \left\{ e^{-.69 \left| Z_i (X_1, \ldots, X_k) \right|} \cos \left[2\pi Z_i (X_1, \ldots, X_k) \right] - 1 \right\}$$

As in the previous case, the Z_i's are linear functions of the X_i's, dependent on the presence or absence of interaction. In addition, the values of the A_i's are determined by the relative activity of the *factors*. In both cases the generated response surfaces have a true (global) optimum response of zero, with all other true responses being negative.

Intuitively, it would be expected that, with respect to Random Search, the adaptive search techniques would perform less well on a multimodal surface with local optima (Figure 4) than on a unimodal surface with only a global optimum (Figure 3). This intuition was supported by the data which was observed in the empirical study. This data is summarized in Section III.C.4.
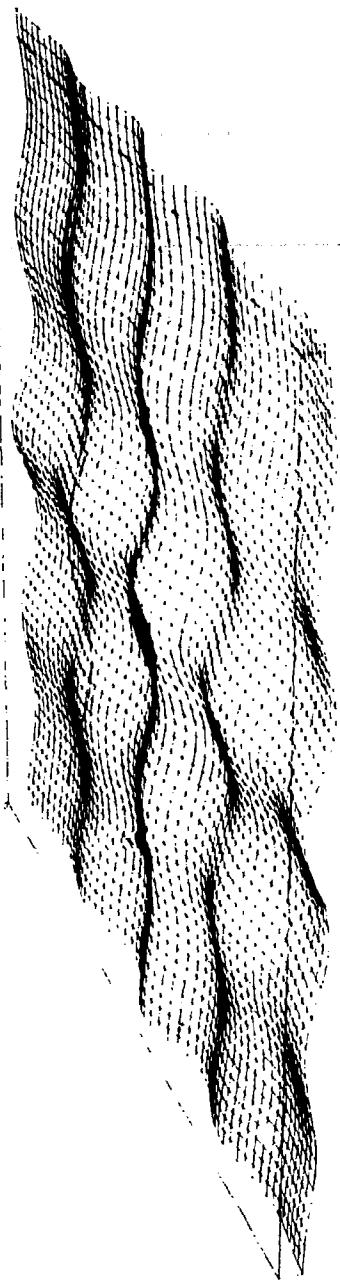
FIG. 3   SURFACE WITH NO LOCAL OPTIMA

AP11000

FIG. 4 SURFACE WITH LOCAL OPTIMA

AP11983

4. Size of Random Error

In this study, two different sizes of random error have been considered, and have arbitrarily been labeled large and small. This section discusses the choice of measurement of the random error.

In attempting to arrive at a suitable method for measuring random error, five alternatives were considered. These were:

(a) Random error proportional to a fixed constant

(b) Random error proportional to the true optimum response

(c) Random error proportional to the true response at the starting point

(d) Random error proportional to the difference between the true optimum response and the true response at the starting point

(e) Random error proportional to the difference between the true optimum response and the true response at some fixed distance from the true optimum point.

In defining the size of random error, it seemed reasonable that the terms "large error" and "small error" be defined consistently so that, for example, small error results in the same error band at all points on a given response surface.

Although alternative (a) does provide a consistent definition of random error, it is completely independent of the response surface itself. If, for example, random error were defined as arising from a probability distribution with standard deviation 10, the effect would be entirely different on a surface having responses in the range (0, .100) than it would be on a surface having responses in the range (0, 100).

At first glance, alternative (b) seems to overcome this situation. However because the initial selection of mathematical equations for generating response surfaces resulted in true optimum responses of zero, this alternative would result in random error which was identically zero. Although the mathematical equations could be modified by the addition of a constant term to provide a non-zero optimum, this procedure would result in objections similar to those raised for alternative (a).

By defining random error proportional to the true response at the starting point, alternative (c) results in inconsistencies. This is due to the fact that, in general, the farther the starting point from the true optimum, the smaller the true response at *that point*. Thus, a starting point relatively close to the true optimum point would tend to result in a smaller error band than would a starting point relatively far from the true optimum. In addition, objections similar to those raised for alternative (b) also exist.

If random error is defined proportional to the difference between the true optimum response and the true response at the starting point, as in alternative (d), a certain amount of stability exists in that the addition of a constant to all responses on the surface does not change the magnitude of the error. However, as in the case of alternative (c), the magnitude of the random error is larger for starting points far from the true optimum point than it is for starting points near to the true optimum point.

Under alternative (e) the size of random error would be consistently defined on any given response surface of given characteristics, regardless of the location of the starting point. In addition, this measure of random error is stable under any translation of the surface in the y-direction. Thus, this alternative was chosen as the definition of random error to be used in the empirical study. Specifically, random error was defined proportional to the difference between the true optimum response and the true response at a point a distance of Δ away (in terms of the coded factors) on the line joining the starting point and the true optimum point. If this difference is denoted by D, small random error was generated from a Normal distribution with mean zero and standard deviation 0.1D, while large random error was generated from a Normal distribution with mean zero and standard deviation 1.0D.

From a practical standpoint, the main disadvantage to this definition of random error is that it may be somewhat difficult for a simulation user to extrapolate it to a simulation context.

5. Distance of Starting Point from True Optimum Point

If, in terms of the original factors, the true optimum point is denoted by $(T_1, \ldots, T_k)$, one aim of the simulation user is to provide a starting point $(C_1, \ldots, C_k)$ that is not extremely distant. To investigate the effect of the distance of starting point from true optimum on search technique performance,

some definition of "near" and "far" distance must be made. Actual distance in the factor space does not provide a suitable measure because different values assigned to a step size Δ_i would result in the distance being traversed in different numbers of steps. It is reasonable to take this into account in order to provide a stable measure.

The user's specification of the C_i's and the Δ_i's result in an initial search region in the locality of the starting point. If the true optimum point is near, there should be a fairly good chance that $C_i - \Delta_i \leq T_i \leq C_i + \Delta_i$ for any i. The chances diminish as the true optimum point becomes more distant. Assume that the user chooses the C_i's and the Δ_i's in such a manner that $\text{Prob}(C_i - \Delta_i \leq T_i \leq C_i + \Delta_i)$ is equal for each i. It is not unrealistic to also assume that each C_i selected by the user is approximately a Normal random variable with expected value T_i and standard deviation which is a function of Δ_i. Definitions of "near" and "far" distance were based on this framework. Specifically, a "near" distance was defined to exist when $\text{Prob}(C_i - \Delta_i \leq T_i \leq C_i + \Delta_i) = 80\%$. Similarly, a "far" distance was defined to exist when $\text{Prob}(C_i - \Delta_i \leq T_i \leq C_i + \Delta_i) = 20\%$. It follows that for near distance, C_i is Normally distributed with mean T_i and standard deviation $0.78\Delta_i$, while for far distance, C_i is Normally distributed with mean T_i and standard deviation $3.95\Delta_i$.

6. Relative Activity of the Controllable Factors

Although a number of controllable factors may be thought to affect the response generated by a computer simulation, in actual practice it is usually found that some of these factors have negligible effect on the response. In other words, the response is insensitive to changes in these factors. Such factors will be labeled inactive, and the remaining factors labeled active.

It is quite conceivable that differences in the relative activity of the controllable factors may dictate different search techniques. Thus, two levels of activity, low and high, were considered. Low activity was defined to exist when only 20% of the controllable factors actually generated the response surface. High activity was defined to exist when 80% of the controllable factors generated the response surface.

7. Presence or Absence of Interaction

Interaction is said to exist when the equation defining the response surface contains terms involving products of the factors. Thus, interaction is absent when no term in the defining equation involves products of the factors. When interaction is present, the optimum value of a given factor depends upon the specific values of other factors. When interaction is absent, the optimum value of a given factor is independent of the specific values other factors may assume. For example, if a response surface were defined as $y = 10 - X_1^2 - X_2^2$ and a maximum were desired, $X_1 = 0$ would provide the optimum response regardless of the value of X_2. However, if the surface were defined as $y = 10 - X_1^2 - 2X_1 X_2 - X_2^2$, the value of X_1 producing the optimum response would depend upon the value of X_2. It can easily be shown, of course, that this value of X_1 would be given by $X_1 = -X_2$.

In Section III.A.1, which deals with the presence or absence of local optima on the response surfaces, the mathematical forms generating the surfaces involved functions of the type $Z_i (X_1, \ldots, X_k)$. For sufaces without interaction, this function, in essence, was defined as $Z_i (X_1, \ldots, X_k) = (X_i - C_i)/\Delta_i$. For surfaces with interaction, $Z_i (X_1, \ldots, X_k)$ was determined by the application of a Helmert transformation to the active factors. This transformation produces various two-factor products in the equation for the response surface. Further details are provided in the following section.

B. CONSTRUCTION OF THE RESPONSE SURFACES

The actual construction of the response surfaces used in the empirical studies was accomplished by a set of computer programs and subprograms designed to provide a response surface according to the desired characteristics (as described in the previous sections), to apply each of the search techniques on the surface, to record the gains obtained by the techniques, and to repeat these steps for a number of iterations on each type of constructed surface. The use of a number of iterations permits a fairly accurate estimate of the expected gain provided by the use of a given search technique on a particular type of response surface.

Before a description is given of the process involved in the construction of a response surface, it should be mentioned that all calculations involved in the construction are based on the coded factors $x_i = (X_i - C_i)/\Delta_i$. As stated

previously, this transformation results in an optimization problem involving factors $(x_1, \ldots, x_k)$, a starting point $(0, \ldots, 0)$, and a step size of one unit for each factor. The remainder of the discussion will refer to the *transformed* factors $(x_1, \ldots, x_k)$, unless the original factors $(X_1, \ldots, X_k)$ are specifically identified.

The process of response surface construction begins with the selection of the point which produces the true optimum. If the true optimum point is denoted, in terms of the original factors, by $(T_1, \ldots, T_k)$, then, in the transformed factor space, the true optimum point has the coordinates given by $t_i = (T_i - C_i)/\Delta_i$, for $i = 1, \ldots, k$. It directly follows from the discussion of the distance criterion[1] that the point $(t_1, \ldots, t_k)$ can be determined by generating each t_i from a Normal distribution with mean 0 and standard deviation which is equal to 3.95 for "far" distance and 0.78 for "near" distance. For each iteration, this is the manner in which the true optimum point is selected.

After this point is selected, the values of the A_i coefficients used in the equation of the response surface (Section III.A.3) are determined. The activity of the factors determines the values of A_i's and the specific factors, x_i, which are to define the response surface. In the computer program, 80% of the factors are selected as active factors in the case of "high" activity and 20% are selected as active in the case of "low" activity. Of the corresponding coefficients, half[2] are set equal to 1.0 and the other half are set equal to 4.0. In each step of this process, the choice of the active x_i's and the selection of the particular values to be given to the A_i's are made randomly.

At this juncture, then, initial preparation for the construction of the surfaces is completed, and there exists a subset of k* controllable factors which are active and which directly affect the response. For convenience in the following discussion, it will be assumed that this subset consists of the first k* of the k factors, i.e., $(x_1, \ldots, x_{k*})$. Of course, the actual k* active factors would probably never be the first k* factors, but having been randomly chosen, would most

---

[1] See Section III.A.5.

[2] It should be noted that if all the A_i's were given the same value, circular contours would result. The way they are defined gives rise to more realistic elliptical contours.

likely be interspersed within the set of k factors. The assumption, however, prevents the mathematical notation from becoming horrendous.

The required mathematical functions Z_i (Section III.A.3) are obtained from equivalent functions z_i of the transformed factors. The z_i, which are defined in terms of the active subset of the controllable factors, are dependent upon the presence or absence of interaction. If interaction is to be absent, each z_i is defined simply as $z_i = x_i - t_i$ which, in terms of the original factors, is equivalent to $Z_i = (X_i - T_i)/\Delta_i$. Thus, the resulting mathematical equation is one which does not contain any terms involving more than one factor. Viewed geometrically, the response surface which is produced in this situation has its axes parallel to the coordinate axes. If interaction is to be present, a Helmert transformation[1] is used to produce the interaction. This transformation is given by:

$$z_1 = \frac{1}{\sqrt{k^*}} \sum_1^{k^*} x_i$$

$$z_j = \frac{1}{\sqrt{j(j-1)}} \left[(j-1)x_j - \sum_1^{j-1} x_i \right] \quad \text{for} \quad 2 \le j \le k^*$$

The resulting equation defining the response surface contains interaction in the form of terms which involve two-factor products. Because of these product terms, the axes of the response surface are not parallel to the coordinate axes.

After the z_i's have been defined, the appropriate choice of surface type is made, dependent upon the presence or absence of local optima. In the absence of local optima, the response surface is generated by

$$y = - \sum_1^{k^*} A_i \left[z_i (x_1, \ldots, x_{k^*}) \right]^2$$

which, regardless of whether or not interaction is present, is a negative definite quadratic equation in $(x_1, \ldots, x_{k^*})$, having a maximum value of $y = 0$. As written, the equation is a canonical representation.

---

[1] See, for example, Kendall and Stuart (1963) for a discussion of this transformation.

If local optima are present, the response surface is generated by

$$y = -\sum_{1}^{k*} A_i \left[e^{-.69 \left| z_i (x_1, \ldots, x_{k*}) \right|} \cos \left[2\pi z_i (x_i, \ldots, x_{k*}) \right] - 1 \right]$$

which also has a maximum value of $y = 0$. This response surface has many local maxima, local minima, and saddle points, as may be seen from Figure 4, which is a surface defined by this equation for the case where only two active controllable factors are involved (i.e., $k* = 2$).

For any particular computer run involving the active controllable factors $(x_1, \ldots, x_{k*})$, the response obtained is calculated from the appropriate equation given above. However, the remaining characteristic, random error, must still be considered. Thus, after a true response has been calculated, random error is generated from the appropriate Normal distribution[1] and is added to this response.

The response surfaces constructed by these procedures were used to evaluate search technique performance.

C. EVALUATION OF SEARCH TECHNIQUE PERFORMANCE

To evaluate the performance of the search technique on response surfaces of differing characteristics, each search technique must be applied to each of the various response surfaces. Furthermore, to provide a reasonably accurate assessment of performance (measured in terms of relative gain), a number of iterations must be made to dampen out random fluctuations arising from the statistical variation present in the procedures used to construct response surfaces of a given type, and from the random error with which the true responses are contaminated. For the relatively small number of characteristics investigated in this study, the number of actual surface searches conducted and, therefore, the number of computer runs used were quite large.

A total of 64 different types of response surfaces were generated, corresponding to all of the combinations of various characteristics considered. These combinations arise from:

---

[1] See Section III.A.4.

Number of controllable factors:	30 or 120
Local Optima:	Presence or Absence
Random Error:	Large or Small
Distance:	Far or Near
Activity:	High or Low
Interaction:	Presence or Absence

Each of the seven search techniques in the prototype "OPTIMIZER" was applied to each type of response surface and, in addition, each was applied for four different values of N, the maximum number of available computer runs. This implies a minimum total of $7 \times 64 \times 4 = 1792$ searches which were conducted. In addition, however, 36 iterations were made for each of these searches, which means that a grand total of 64,512 searches were made in this investigation[1]

1. Definition of the Overall Search Region

Because the starting point of $(0, \ldots, 0)$ and the initial step sizes of unity remained constant for each search, the actual application of the search techniques could, except in the case of Random Search, proceed directly without requiring any additional data or computation. Random Search, however, required the assignment of an overall search region. Because the size of the overall region directly affects the performance of Random Search, this assignment posed a major problem.

Since it was thought that Random Search would make a relatively poor showing on the unimodal surfaces, it was decided to lean over backward, so to speak, to define an overall region which provided an opportunity for Random Search to perform well. To this end, the overall search region was constructed in such a manner that it always contained the true optimum point within a relatively small volume. Specifically, if t_j, the jth coordinate of the optimum point, fell in the interval $(-.75, .75)$ the overall region for x_j was defined as $(-1, 1)$. If $t_j \geq .75$, the overall region for x_j was defined as $(-1, t_j + .25)$, and if $t_j \leq -.75$, the overall region for x_j was defined as $(t_j - .25, 1)$.

It was reasoned that if Random Search made a poor showing under these somewhat ideal conditions, this would build a strong case for jettisoning the technique. Such a poor showing was anticipated, but did not materialize.[1]

2. Procedure

For each combination of characteristics, the applicable computer routines generated the response surface corresponding to the characteristics, calculated the true response at the starting point, applied each of the seven search techniques to the surface, recorded the gain achieved by each technique for each of the four values of N, and then repeated these steps until a total of 36 iterations were completed. In essence, the overall investigation could be regarded as being based on a 2^6 x 4 x 7 factorial experiment with 36 observations per cell.

Different sequences of pseudo-random numbers were used for generating different response surfaces within a given type of surface defined by a certain combination of characteristics. Although a given type of surface would be characterized by, say, "near" distance and "low" activity, surfaces of this type, in general, would have starting points with different coordinates (although still conforming to the "near" distance criterion) and would consist of different subsets of active factors (although the same precentage of factors would be active). Different sequences of pseudo-random numbers were also used in the generation of the random error terms which were added to the true responses to yield the observed responses.

3. Preliminary Data Analysis

The procedure outlined above provides for the statistical independence of the responses obtained from the iterations of a given search technique on a specific type[2] of surface, as well as the independence of those obtained from different types of surfaces. However, because each of the seven search techniques was applied to every surface of a given type, the responses obtained by

---

[1] See the discussion in Section III. D.

[2] In this discussion a surface type implies a surface determined by a set of characteristics. Surfaces corresponding to different sets of characteristics are thus of different types.

the different search techniques on a specific type of surface are correlated. The resulting lack of statistical independence in the overall experiment creates difficulty in an analysis of the data, because statistical independence is an assumption on which most standard methods of statistical analysis are based. However, certain advantages offset this disadvantage. One major advantage is that by using iterations in which the search techniques were applied to the same response surfaces within a type, differences between the observed responses for each technique were not contaminated by any effect due to differences in the individual response surfaces[1]. Another advantage to using the same surfaces for iterations of each of the seven search techniques is that computer time is saved by having to generate only 36 surfaces of a certain type instead of 36 x 7 = 252 surfaces.

Despite the lack of statistical independence, the manner in which the experiment was conducted still permits an analysis of variance (ANOVA) to be performed, with such an analysis based on the assumption of nesting[2] or repeated measures[3]. However, ANOVA seems inappropriate for analyzing the experimental data because it does not provide a direct answer to the previously asked question about the best technique to use on a surface of given characteristics and the gain to be expected from using this technique. In essence, a more desirable outcome of the empirical studies is a set of payoff matrices on which to base the selection of a search technique.

A preliminary data analysis was adopted in an attempt to reduce the accumulated data by determining which of the response surface characteristics had an effect on the gains provided by a given search technique. This preliminary data analysis involved an ANOVA for each search technique and combination of k and N. A standard ANOVA based on the assumption of a completely randomized factorial experiment could be performed in each case because of the independence inherent in each combination. As one would expect, correlation did, however, exist between various of the ANOVA. Nonetheless, each ANOVA strongly indicated that the gains provided by the search techniques were quite

---

[1] A discussion of the pros and cons of a situation like this is given in Hillier and Lieberman (1967).

[2] See, for example, Scheffé (1959).

[3] See, for example, Winer (1962).

sensitive to all the characteristics considered, except for the presence or absence of interaction. This information indicates that all the characteristics (except interaction) must be considered in the construction of any payoff matrix.

4. The Payoff Matrices

The simulation user who wishes to use a search technique in an attempt to locate an optimum solution essentially finds himself in a decision theory framework[1], playing a game against Nature. In decision theory terminology, the possible states of Nature correspond to the particular combinations of characteristics possessed by the response surface which is generated by the simulation. The user has a set of possible actions corresponding to each of the seven search techniques. For each action-state of Nature combination there is a payoff, which in this situation may be assumed to be expected gain.

Before the user can make any rational decision as to the choice of an action (i.e., a search technique), however, the value of the various payoffs (i.e., expected gains) must be specified. Although the true payoffs are unknown, the empirical studies have provided relatively good estimates of these payoffs. Figures 5 through 12 contain the payoff matrices based on the estimated expected gains. The standard error of each estimated expected gain is given to indicate the accuracy of the estimation.

If a user were faced with a simulation involving 30 controllable factors and had available a maximum of 45 computer runs, he could base his selection of a search technique on the data contained in the payoff matrix given in Figure 7. For example, suppose the state of Nature was known to produce a response surface with:

        (a) no local optima

        (b) a true optimum far from the starting point

        (c) large random error

        (d) high factor activity.

---

[1] For a discussion of decision theory see Chernoff and Moses (1959) or Luce and Raiffa (1957), for example.

LOCAL OPTIMA	DISTANCE	ERROR	ACTIVITY	RANDOM SEARCH	SINGLE-FACTOR	SINGLE-FACTOR (ACCEL)	RSM-I	RSM-I (ACCEL)	RSM-II	RSM-II (ACCEL)
ABSENT	NEAR	SMALL	LOW	0.474 (0.033)	0.194 (0.028)	0.177 (0.027)	0.207 (0.032)	0.207 (0.032)	0.145 (0.022)	0.140 (0.022)
ABSENT	NEAR	SMALL	HIGH	0.230 (0.022)	0.116 (0.016)	0.097 (0.017)	0.003 (0.002)	0.031 (0.002)	0.118 (0.011)	0.115 (0.010)
ABSENT	NEAR	LARGE	LOW	0.351 (0.038)	0.040 (0.019)	0.029 (0.013)	0.091 (0.034)	0.069 (0.034)	0.073 (0.021)	0.032 (0.016)
ABSENT	NEAR	LARGE	HIGH	0.028 (0.030)	-0.047 (0.017)	-0.037 (0.017)	-0.175 (0.031)	-0.086 (0.033)	-0.034 (0.013)	-0.020 (0.015)
ABSENT	FAR	SMALL	LOW	0.466 (0.009)	0.135 (0.018)	0.139 (0.020)	0.328 (0.013)	0.328 (0.013)	0.108 (0.013)	0.108 (0.013)
ABSENT	FAR	SMALL	HIGH	0.744 (0.007)	0.112 (0.012)	0.112 (0.010)	0.136 (0.006)	0.136 (0.006)	0.052 (0.002)	0.052 (0.002)
ABSENT	FAR	LARGE	LOW	0.869 (0.009)	0.136 (0.027)	0.173 (0.023)	0.307 (0.013)	0.304 (0.012)	0.102 (0.010)	0.103 (0.009)
ABSENT	FAR	LARGE	HIGH	0.750 (0.008)	0.090 (0.011)	0.084 (0.011)	0.140 (0.006)	0.137 (0.006)	0.043 (0.003)	0.044 (0.003)
PRESENT	NEAR	SMALL	LOW	0.368 (0.024)	0.094 (0.017)	0.047 (0.018)	0.145 (0.020)	0.157 (0.020)	0.107 (0.014)	0.117 (0.015)
PRESENT	NEAR	SMALL	HIGH	0.162 (0.016)	0.076 (0.007)	0.028 (0.007)	0.052 (0.011)	0.061 (0.010)	0.017 (0.005)	0.023 (0.005)
PRESENT	NEAR	LARGE	LOW	0.044 (0.045)	0.018 (0.014)	0.021 (0.016)	-0.037 (0.044)	-0.124 (0.044)	-0.007 (0.009)	0.001 (0.015)
PRESENT	NEAR	LARGE	HIGH	0.017 (0.019)	0.001 (0.005)	0.008 (0.008)	-0.017 (0.017)	0.003 (0.009)	0.007 (0.006)	-0.001 (0.001)
PRESENT	FAR	SMALL	LOW	0.247 (0.019)	0.023 (0.008)	0.011 (0.006)	0.043 (0.012)	0.075 (0.011)	0.053 (0.006)	0.047 (0.006)
PRESENT	FAR	SMALL	HIGH	0.116 (0.010)	0.007 (0.002)	0.007 (0.002)	0.022 (0.006)	0.024 (0.006)	0.011 (0.003)	0.005 (0.003)
PRESENT	FAR	LARGE	LOW	0.031 (0.027)	-0.002 (0.005)	0.013 (0.006)	0.010 (0.016)	-0.001 (0.015)	0.006 (0.003)	0.010 (0.007)
PRESENT	FAR	LARGE	HIGH	0.043 (0.013)	-0.000 (0.003)	0.002 (0.004)	0.003 (0.006)	0.006 (0.005)	0.202 (0.003)	-0.000 (0.002)

FIG. 5 PAYOFF MATRIX FOR X-30, N-15 SHOWING ESTIMATED EXPECTED GAINS AND ASSOCIATED STANDARD ERRORS IN PARENTHESES

LOCAL OPTIMA	DISTANCE	ERROR	ACTIVITY	RANDOM SEARCH	SINGLE-FACTOR	SINGLE-FACTOR (ACCEL)	RSM-I	RSM-I (ACCEL)	RSM-II	RSM-II (ACCEL)
ABSENT	NEAR	SMALL	LOW	0.596 (0.031)	0.361 (0.033)	0.357 (0.033)	0.256 (0.034)	0.255 (0.034)	0.342 (0.024)	0.339 (0.024)
ABSENT	NEAR	SMALL	HIGH	0.313 (0.021)	0.273 (0.019)	0.260 (0.019)	0.009 (0.004)	0.006 (0.004)	0.199 (0.008)	0.199 (0.008)
ABSENT	NEAR	LARGE	LOW	0.392 (0.043)	0.122 (0.029)	0.106 (0.027)	0.161 (0.044)	0.157 (0.045)	0.129 (0.027)	0.074 (0.024)
ABSENT	NEAR	LARGE	HIGH	0.021 (0.033)	-0.045 (0.021)	-0.044 (0.019)	-0.254 (0.058)	-0.142 (0.032)	-0.023 (0.018)	-0.010 (0.016)
ABSENT	FAR	SMALL	LOW	0.907 (0.006)	0.350 (0.010)	0.331 (0.031)	0.386 (0.010)	0.386 (0.010)	0.183 (0.010)	0.183 (0.010)
ABSENT	FAR	SMALL	HIGH	0.783 (0.026)	0.239 (0.014)	0.251 (0.016)	0.171 (0.005)	0.171 (0.005)	0.069 (0.002)	0.049 (0.002)
ABSENT	FAR	LARGE	LOW	0.01 (0.008)	0.342 (0.031)	0.310 (0.030)	0.350 (0.010)	0.350 (0.010)	0.161 (0.009)	0.163 (0.008)
ABSENT	FAR	LARGE	HIGH	0.785 (0.006)	0.200 (0.016)	0.197 (0.014)	0.178 (0.005)	0.178 (0.005)	0.061 (0.003)	0.065 (0.002)
PRESENT	NEAR	SMALL	LOW	0.419 (0.022)	0.116 (0.017)	0.124 (0.017)	0.167 (0.020)	0.185 (0.019)	0.218 (0.021)	0.240 (0.020)
PRESENT	NEAR	SMALL	HIGH	0.197 (0.014)	0.033 (0.007)	0.038 (0.008)	0.067 (0.011)	0.066 (0.011)	0.040 (0.007)	0.061 (0.007)
PRESENT	NEAR	LARGE	LOW	0.022 (0.047)	0.031 (0.015)	0.021 (0.015)	-0.035 (0.040)	0.021 (0.019)	0.015 (0.013)	0.010 (0.015)
PRESENT	NEAR	LARGE	HIGH	0.028 (0.019)	0.005 (0.006)	0.003 (0.008)	-0.008 (0.011)	0.013 (0.009)	0.001 (0.006)	-0.057 (0.047)
PRESENT	FAR	SMALL	LOW	0.293 (0.017)	0.037	0.030 (0.004)	0.081 (0.012)	0.093 (0.015)	0.064 (0.013)	0.071 (0.013)
PRESENT	FAR	SMALL	HIGH	0.127 (0.010)	0.009 (0.002)	0.010 (0.003)	0.024 (0.005)	0.029 (0.006)	0.013 (0.004)	0.013 (0.004)
PRESENT	FAR	LARGE	LOW	0.057 (0.024)	-0.002 (0.001)	-0.014 (0.004)	0.019 (0.016)	0.327 (0.015)	0.026 (0.009)	0.009 (0.010)
PRESENT	FAR	LARGE	HIGH	0.039 (0.012)	-0.001 (0.003)	0.005 (0.004)	0.011 (0.077)	0.009 (0.005)	0.003 (0.005)	0.001 (0.003)

FIG. 6 PAYOFF MATRIX FOR k=30, M=33 SHOWING ESTIMATED EXPECTED GAINS AND ASSOCIATED STANDARD ERRORS IN PARENTHESES

LOCAL OPTIMA	DISTANCE	ERROR	ACTIVITY	RANDOM SEARCH	SINGLE-FACTOR	SINGLE-FACTOR (ACCEL)	RSM-I	PSM-I (ACCEL)	RSM-II	RSM-II (ACCEL)
ABSENT	NEAR	SMALL	LOW	0.647 (0.030)	0.494 (0.033)	0.487 (0.033)	0.522 (0.034)	0.515 (0.034)	0.786 (0.024)	0.778 (0.025)
ABSENT	NEAR	SMALL	HIGH	0.346 (0.006)	0.374 (0.020)	0.355 (0.021)	0.045 (0.030)	0.098 (0.031)	0.762 (0.017)	0.768 (0.010)
ABSENT	NEAR	LARGE	LOW	0.446 (0.042)	0.158 (0.030)	0.172 (0.029)	0.444 (0.046)	0.459 (0.045)	0.231 (0.036)	0.216 (0.033)
ABSENT	NEAR	LARGE	HIGH	0.609 (0.042)	-0.041 (0.021)	-0.045 (0.014)	0.167 (0.046)	0.145 (0.045)	-0.010 (0.020)	0.043
ABSENT	FAR	SMALL	LOW	0.923 (0.034)	0.507	0.494 (0.034)	0.944 (0.038)	0.945 (0.033)	0.939 (0.020)	0.940 (0.021)
ABSENT	FAR	SMALL	HIGH	0.606 (0.006)	0.232 (0.032)	0.034 (0.034)	0.006 (0.006)	0.005 (0.005)	0.007 (0.007)	0.007 (0.007)
ABSENT	FAR	LARGE	LOW	0.798 (0.006)	0.323 (0.017)	0.343	0.921	0.321 (0.005)	0.915 (0.004)	0.916 (0.005)
ABSENT	FAR	LARGE	HIGH	0.915 (0.006)	0.469 (0.037)	0.426 (0.019)	0.906 (0.003)	0.913	0.884 (0.012)	0.879 (0.012)
ABSENT	FAR	LARGE	HIGH	0.707 (0.006)	0.265 (0.016)	0.259 (0.017)	0.910 (0.005)	0.895 (0.006)	0.780 (0.009)	0.772 (0.009)
PRESENT	NEAR	SMALL	LOW	0.446 (0.019)	0.144 (0.018)	0.154 (0.017)	0.251 (0.021)	0.264 (0.022)	0.293 (0.022)	0.315 (0.022)
PRESENT	NEAR	SMALL	HIGH	0.177 (0.014)	0.038 (0.007)	0.038 (0.008)	0.107 (0.014)	0.090 (0.013)	0.075 (0.012)	0.097 (0.011)
PRESENT	NEAR	LARGE	LOW	0.004 (0.050)	0.075 (0.015)	0.043 (0.017)	-0.013 (0.043)	0.016 (0.023)	0.018 (0.016)	0.021
PRESENT	NEAR	LARGE	HIGH	0.035 (0.019)	0.006	0.005 (0.009)	-0.016 (0.012)	0.009 (0.009)	-0.004 (0.007)	-0.001 (0.009)
PRESENT	FAR	SMALL	LOW	0.322 (0.015)	0.045 (0.011)	0.037	0.118 (0.015)	0.110 (0.012)	0.103 (0.015)	0.073
PRESENT	FAR	SMALL	HIGH	0.130 (0.010)	0.009 (0.033)	0.011 (0.004)	0.040 (0.007)	0.040 (0.006)	0.020 (0.005)	0.020 (0.005)
PRESENT	FAR	LARGE	LOW	0.077 (0.027)	-0.009	0.013 (0.004)	0.011 (0.018)	0.228 (0.015)	0.227 (0.009)	0.010 (0.010)
PRESENT	FAR	LARGE	HIGH	0.046 (0.012)	-0.001 (0.004)	0.004 (0.004)	0.009 (0.007)	0.010 (0.006)	0.003 (0.006)	-0.002 (0.004)

FIG. 7 PAYOFF MATRIX FOR K=30, N=45 SHOWING ESTIMATED EXPECTED GAINS AND ASSOCIATED STANDARD ERRORS IN PARENTHESES

LOCAL OPTIMA	DISTANCE	FREQ	ACTIVITY	RANDOM SEARCH	SINGLE-FACTOR	SINGLE-FACTOR (ACCEL)	PSM-I	PSM-I (ACCEL)	RSM-II	RSM-II (ACCEL)
ABSENT	NEAR	SMALL	LOW	0.669 (0.029)	0.621 (0.032)	0.606 (0.033)	0.532 (0.055)	0.514 (0.054)	0.782 (0.023)	0.769 (0.025)
ABSENT	NEAR	SMALL	HIGH	0.581 (0.019)	0.507 (0.019)	0.478 (0.021)	0.086 (0.031)	0.395 (0.030)	0.791 (0.015)	0.816 (0.001)
ABSENT	NEAR	LARGE	LOW	0.479 (0.037)	0.170 (0.035)	0.161 (0.031)	0.426 (0.065)	0.479 (0.045)	0.247 (0.034)	0.208 (0.001)
ABSENT	NEAR	LARGE	HIGH	0.079 (0.034)	-0.038 (0.035)	-0.037 (0.031)	0.136 (0.033)	0.146 (0.033)	-0.010 (0.021)	0.037 (0.021)
ABSENT	FAR	SMALL	LOW	0.934 (0.005)	0.706 (0.020)	0.668 (0.019)	0.976 (0.003)	0.977 (0.003)	0.952 (0.021)	0.952 (0.001)
ABSENT	FAR	SMALL	HIGH	0.805 (0.005)	0.442 (0.018)	0.455 (0.019)	0.950 (0.002)	0.954 (0.002)	0.931 (0.004)	0.932 (0.004)
ABSENT	FAR	LARGE	LOW	0.625 (0.006)	0.651 (0.016)	0.667 (0.027)	0.949 (0.002)	0.954 (0.004)	0.893 (0.004)	0.890 (0.004)
ABSENT	FAR	LARGE	HIGH	0.805 (0.006)	0.369 (0.016)	0.349 (0.019)	0.911 (0.004)	0.904 (0.005)	0.785 (0.010)	0.776 (0.009)
PRESENT	NEAR	SMALL	LOW	0.478 (0.020)	0.172 (0.017)	0.179 (0.017)	0.248 (0.021)	0.265 (0.020)	0.309 (0.099)	0.357 (0.020)
PRESENT	NEAR	SMALL	HIGH	0.198 (0.013)	0.037 (0.007)	0.045 (0.008)	0.104 (0.013)	0.100 (0.013)	0.084 (0.012)	0.095 (0.011)
PRESENT	NEAR	LARGE	LOW	-0.030 (0.057)	0.023 (0.016)	0.045 (0.017)	-0.030 (0.013)	0.058 (0.028)	0.029 (0.017)	0.028 (0.018)
PRESENT	NEAR	LARGE	HIGH	0.043 (0.016)	0.006 (0.016)	0.006 (0.009)	-0.002 (0.044)	0.017 (0.011)	0.005 (0.017)	0.009 (0.009)
PRESENT	FAR	SMALL	LOW	0.348 (0.016)	0.056 (0.011)	0.043 (0.011)	0.113 (0.013)	0.112 (0.011)	0.104 (0.007)	0.111 (0.017)
PRESENT	FAR	SMALL	HIGH	0.134 (0.016)	0.010 (0.003)	0.011 (0.003)	0.045 (0.014)	0.046 (0.014)	0.227 (0.017)	0.025 (0.017)
PRESENT	FAR	LARGE	LOW	0.071 (0.027)	-0.000 (0.005)	-0.012 (0.003)	0.012 (0.007)	0.030 (0.007)	0.035 (0.005)	0.013 (0.006)
PRESENT	FAR	LARGE	HIGH	0.049 (0.012)	-0.001 (0.006)	0.004 (0.005)	0.014 (0.007)	0.011 (0.008)	0.005 (0.006)	0.000 (0.003)

FIG. 8 PAYOFF MATRIX FOR k=30, N=60 SHOWING ESTIMATED EXPECTED GAINS AND ASSOCIATED STANDARD ERRORS IN PARENTHESES

FIG. 9 PAYOFF MATRIX FOR k=128, N=88 SHOWING ESTIMATED EXPECTED GAINS AND ASSOCIATED STANDARD ERRORS IN PARENTHESES

LOCAL OPTIMA	DISTANCE	ERROR	ACTIVITY	RANDOM SEARCH	SINGLE-FACTOR	SINGLE-FACTOR (ACCEL)	RSM-I	RSM-I (ACCEL)	RSM-II	RSM-II (ACCEL)
ABSENT	NEAR	SMALL	LOW	0.334 (0.022)	0.133 (0.015)	0.175 (0.015)	0.007 (0.035)	0.008 (0.005)	0.094 (0.012)	0.102 (0.012)
ABSENT	NEAR	SMALL	HIGH	0.057 (0.011)	0.023 (0.005)	0.015 (0.036)	0.0	0.0	0.029 (0.004)	0.029 (0.004)
ABSENT	NFAR	LARGE	LOW	0.111 (0.036)	-0.005 (0.007)	-0.095 (0.013)	-0.276 (0.051)	-0.215 (0.040)	0.006 (0.002)	0.002 (0.003)
ABSENT	NEAR	LARGE	HIGH	-0.157 (0.021)	-0.342 (0.006)	-0.056 (0.014)	-0.219 (0.050)	-0.155 (0.044)	-0.016 (0.003)	-0.013 (0.004)
ABSENT	FAR	SMALL	LOW	0.807 (0.005)	0.184 (0.014)	0.181 (0.014)	0.195 (0.005)	0.195 (0.005)	0.052 (0.003)	0.052 (0.003)
ABSENT	FAR	SMALL	HIGH	0.604 (0.003)	0.121 (0.007)	0.119 (0.007)	-0.066 (0.002)	0.066 (0.002)	0.019 (0.001)	0.019 (0.001)
ABSENT	FAR	LARGE	LOW	0.802 (0.006)	0.102 (0.014)	0.104 (0.013)	0.198 (0.006)	0.198 (0.006)	0.043 (0.004)	0.043 (0.004)
ABSENT	FAR	LARGE	HIGH	0.686 (0.004)	0.014 (0.002)	0.017 (0.003)	0.066 (0.006)	0.366 (0.004)	0.009 (0.001)	0.009 (0.001)
PRESENT	NFAR	SMALL	LOW	0.220 (0.013)	0.015 (0.008)	0.011 (0.005)	0.076 (0.004)	0.061 (0.012)	0.012 (0.003)	0.013 (0.004)
PRESENT	NFAR	SMALL	HIGH	0.079 (0.006)	0.001 (0.001)	-0.001	0.001 (0.005)	-0.000	0.001 (0.001)	0.001
PRESENT	NEAR	LARGE	LOW	0.016 (0.020)	-0.006 (0.003)	-0.002 (0.002)	-0.005 (0.013)	-0.004 (0.002)	-0.001 (0.002)	0.002 (0.003)
PRESENT	NFAR	LARGE	HIGH	-0.005 (0.011)	-0.001 (0.003)	-0.000 (0.001)	-0.016 (0.021)	-0.008 (0.001)	0.004 (0.001)	-0.005 (0.001)
PRESENT	FAR	SMALL	LOW	0.138 (0.006)	0.006 (0.001)	0.005 (0.004)	-0.004 (0.004)	-0.021 (0.004)	0.001 (0.001)	0.003 (0.002)
PRESENT	FAR	SMALL	HIGH	0.058 (0.011)	-0.006 (0.003)	-0.004 (0.002)	0.035 (0.006)	0.037 (0.007)	0.003 (0.002)	0.002 (0.001)
PRESENT	FAR	LARGE	LOW	0.049 (0.012)	0.002 (0.002)	0.003 (0.001)	0.012 (0.028)	0.007 (0.006)	-0.002 (0.001)	-0.000 (0.001)
PRESENT	FAR	LARGE	HIGH	0.034 (0.006)	0.001 (0.001)	0.000 (0.001)	-0.007 (0.053)	-0.003 (0.002)	-0.000 (0.001)	-0.001 (0.001)

-309-

LOCAL OPTIMA	DISTANCE	FREQW	ACTIVITY	RANDOM SEARCH	SINGLE-FACTOR	SINGLE-FACTOR (ACCEL)	RSM-1	PSM-1 (ACCEL)	RSM-11	RSM-11 (ACCEL)
ABSENT	NEAR	SMALL	LOW	0.415 (0.021)	0.790 (0.074)	0.257 (0.022)	0.145 (0.036)	0.113 (0.033)	0.511 (0.017)	0.537 (0.014)
ABSENT	NEAR	SMALL	HIGH	0.088 (0.012)	0.054 (0.008)	0.045 (0.004)	0.0 (0.0)	0.0 (0.0)	0.473 (0.008)	0.490 (0.010)
ABSENT	NEAR	LARGE	LOW	0.156 (0.034)	-0.002 (0.036)	-0.004 (0.014)	0.353 (0.047)	0.345 (0.046)	0.004 (0.003)	0.003 (0.005)
ABSENT	NEAR	LARGE	HIGH	-0.155 (0.025)	-0.050 (0.010)	-0.064 (0.014)	0.112 (0.020)	0.085 (0.020)	-0.015 (0.007)	-0.000 (0.002)
ABSENT	FAR	SMALL	LOW	0.830 (0.024)	0.408 (0.018)	0.394 (0.014)	0.880 (0.005)	0.880 (0.005)	0.910 (0.004)	0.910 (0.004)
ABSENT	FAR	SMALL	HIGH	0.763 (0.003)	0.258 (0.008)	0.253 (0.008)	0.617 (0.004)	0.616 (0.004)	0.893 (0.002)	0.894 (0.002)
ABSENT	FAR	LARGE	LOW	0.828 (0.005)	0.206 (0.019)	0.203 (0.017)	0.877 (0.005)	0.877 (0.005)	0.775 (0.008)	0.769 (0.007)
ABSENT	FAR	LARGE	HIGH	0.705 (0.004)	0.018 (0.003)	0.021 (0.003)	0.589 (0.004)	0.590 (0.004)	0.241 (0.009)	0.233 (0.010)
ABSENT	FAR	SMALL	LOW	0.233 (0.015)	0.018 (0.008)	0.016 (0.006)	0.085 (0.012)	0.078 (0.012)	0.040 (0.009)	0.032 (0.007)
PRESENT	NEAR	SMALL	HIGH	0.085 (0.009)	0.002 (0.001)	0.000 (0.000)	-0.000 (0.004)	0.006 (0.004)	0.002 (0.002)	0.005 (0.003)
PRESENT	NEAR	LARGE	LOW	0.013 (0.021)	0.001 (0.003)	-0.004 (0.005)	-0.011 (0.013)	-0.010 (0.010)	0.004 (0.005)	0.009 (0.005)
PRESENT	NEAR	LARGE	HIGH	-0.002 (0.011)	-0.001 (0.003)	0.000 (0.001)	-0.016 (0.004)	-0.017 (0.004)	0.001 (0.001)	-0.003 (0.003)
PRESENT	FAR	SMALL	LOW	0.168 (0.011)	-0.006 (0.003)	-0.001 (0.001)	0.045 (0.004)	0.045 (0.004)	0.008 (0.003)	0.007 (0.004)
PRESENT	FAR	SMALL	HIGH	0.051 (0.005)	0.000 (0.001)	0.001 (0.001)	0.098 (0.003)	0.006 (0.002)	0.002 (0.001)	0.001 (0.001)
PRESENT	FAR	LARGE	LOW	0.048 (0.013)	0.004 (0.003)	-0.004 (0.001)	0.003 (0.002)	0.010 (0.002)	-0.001 (0.001)	0.002 (0.002)
PRESENT	FAR	LARGE	HIGH	0.023 (0.006)	0.001 (0.001)	0.000 (0.001)	-0.003 (0.003)	-0.001 (0.002)	0.001 (0.001)	-0.001 (0.001)

FIG. 10   PAYOFF MATRIX FOR k=120, N=132 SHOWING ESTIMATED EXPECTED GAINS AND ASSOCIATED STANDARD ERRORS IN PARENTHESES

LOCAL OPTIMA	DISTANCE	SPEED	ACTIVITY	RANDOM SEARCH	SINGLE-FACTOR	SINGLE-FACTOR (ACCEL)	RSM-I	RSM-I (ACCEL)	RSM-II	RSM-II (ACCEL)
ABSENT	NEAR	SMALL	LOW	0.442 (0.070)	0.390 (0.021)	0.388 (0.021)	0.151 (0.038)	0.115 (0.034)	0.698 (0.011)	0.700 (0.009)
ABSENT	NEAR	SMALL	HIGH	0.101 (0.011)	0.066 (0.010)	0.057 (0.009)	0.0 (0.0)	0.0 (0.0)	0.480 (0.008)	0.488 (0.010)
ABSENT	NEAR	LARGE	LOW	0.162 (0.028)	-0.002 (0.026)	-0.008 (0.014)	0.284 (0.040)	0.260 (0.036)	0.009 (0.008)	0.010 (0.005)
ABSENT	NEAR	LARGE	HIGH	-0.134 (0.023)	-0.056 (0.010)	-0.069 (0.014)	0.092 (0.017)	0.059 (0.017)	-0.018 (0.004)	-0.011 (0.004)
ABSENT	FAR	SMALL	LOW	0.837 (0.064)	0.544 (0.021)	0.527 (0.021)	0.963 (0.003)	0.967 (0.003)	0.936 (0.003)	0.936 (0.003)
ABSENT	FAR	SMALL	HIGH	0.769 (0.003)	0.349 (0.008)	0.343 (0.008)	0.929 (0.002)	0.928 (0.002)	0.900 (0.002)	0.901 (0.002)
ABSENT	FAR	LARGE	LOW	0.838 (0.005)	0.285 (0.019)	0.280 (0.019)	0.920 (0.004)	0.923 (0.003)	0.778 (0.008)	0.774 (0.007)
ABSENT	FAR	LARGE	HIGH	0.710 (0.004)	0.019 (0.003)	0.023 (0.004)	0.874 (0.003)	0.872 (0.003)	0.244 (0.009)	0.235 (0.011)
PRESENT	NEAR	SMALL	LOW	0.254 (0.014)	0.021 (0.008)	0.018 (0.007)	0.096 (0.012)	0.092 (0.013)	0.044 (0.010)	0.036 (0.007)
PRESENT	NEAR	SMALL	HIGH	0.095 (0.009)	0.007 (0.011)	0.000 (0.002)	0.007 (0.005)	0.013 (0.004)	0.002 (0.002)	0.004 (0.003)
PRESENT	NEAR	LARGE	LOW	0.027 (0.021)	0.001 (0.003)	-0.006 (0.005)	-0.015 (0.013)	-0.004 (0.011)	0.002 (0.005)	0.006 (0.005)
PRESENT	NEAR	LARGE	HIGH	-0.002 (0.011)	-0.001 (0.021)	0.000 (0.005)	-0.017 (0.013)	-0.014 (0.011)	0.001 (0.003)	-0.003 (0.003)
PRESENT	FAR	SMALL	LOW	0.174 (0.012)	0.008 (0.003)	0.037 (0.004)	0.049 (0.003)	0.054 (0.008)	0.010 (0.003)	0.010 (0.005)
PRESENT	FAR	SMALL	HIGH	0.062 (0.005)	0.006 (0.001)	-0.001 (0.001)	-0.009 (0.003)	0.010 (0.003)	-0.002 (0.002)	0.002 (0.002)
PRESENT	FAR	LARGE	LOW	0.050 (0.012)	0.003 (0.003)	0.004 (0.002)	0.013 (0.008)	0.014 (0.007)	-0.002 (0.002)	0.003 (0.002)
PRESENT	FAR	LARGE	HIGH	0.028 (0.006)	0.000 (0.001)	-0.000 (0.001)	-0.001 (0.008)	-0.004 (0.003)	0.001 (0.001)	-0.000 (0.001)

FIG. 11 PAYOFF MATRIX FOR k=128, N=180 SHOWING ESTIMATED EXPECTED GAINS AND ASSOCIATED STANDARD ERRORS IN PARENTHESES

CRITICAL OPTIMA	DISTANCE	FRKND	ACTIVITY	RANDOM SEARCH	SINGLE-FACTOR	SINGLE-FACTOR (ACCFL)	RSM-I	RSM-I (ACCEL)	RSM-II	RSM-II (ACCEL)
ABSENT	NEAR	SMALL	LOW	0.467 (0.019)	0.518 (0.020)	0.512 (0.013)	0.153 (0.038)	0.120 (0.035)	0.726 (0.010)	0.715 (0.011)
ABSENT	NEAR	SMALL	HIGH	0.111 (0.014)	0.092 (0.011)	0.078 (0.010)	0.0	0.0	0.493 (0.008)	0.499 (0.011)
ABSENT	NEAR	LARGE	LOW	0.178 (0.032)	-0.000 (0.007)	-0.001	0.255 (0.036)	0.271 (0.037)	0.006 (0.008)	0.005 (0.006)
ABSENT	NEAR	LARGE	HIGH	-0.114 (0.021)	-0.051 (0.010)	-0.074 (0.015)	0.066 (0.013)	0.031 (0.015)	-0.018 (0.004)	-0.014 (0.004)
ABSENT	FAR	SMALL	LOW	0.845 (0.004)	0.708 (0.018)	0.675 (0.019)	0.984 (0.013)	0.981 (0.002)	0.942 (0.004)	0.941 (0.003)
ABSENT	FAR	SMALL	HIGH	0.715 (0.003)	0.470 (0.009)	0.457 (0.009)	0.948 (0.003)	0.948 (0.003)	0.906 (0.002)	0.908 (0.002)
ABSENT	FAR	LARGE	LOW	0.849 (0.004)	0.385 (0.022)	0.351 (0.022)	0.928 (0.004)	0.930 (0.003)	0.781 (0.009)	0.775 (0.007)
ABSENT	FAR	LARGE	HIGH	0.717 (0.004)	0.024 (0.003)	0.027 (0.004)	0.873 (0.003)	0.872 (0.014)	0.244 (0.009)	0.234
PRESENT	NEAR	SMALL	LOW	0.257 (0.015)	0.074 (0.001)	0.020 (0.002)	0.104 (0.003)	0.103	0.047 (0.010)	0.047
PRESENT	NEAR	SMALL	HIGH	0.092	0.002	0.301	0.015 (0.007)	0.014 (0.005)	0.007 (0.003)	0.007
PRESENT	NEAR	LARGE	LOW	0.009 (0.00P)	0.001 (0.002)	0.002	-0.013	-0.007	0.005	0.003
PRESENT	NEAR	LARGE	HIGH	0.033 (0.021)	0.003 (0.005)	0.006	-0.015 (0.010)	-0.010 (0.006)	0.002	0.003
PRESENT	FAR	SMALL	LOW	-0.006	-0.001 (0.001)	0.000	-0.015 (0.006)	-0.014	0.002 (0.001)	0.003
PRESENT	FAR	SMALL	HIGH	0.011 (0.011)	0.008 (0.003)	0.008	0.059 (0.008)	0.055	0.010 (0.003)	0.011
PRESENT	FAR	LARGE	LOW	0.067 (0.005)	0.060 (0.001)	0.001	0.010 (0.003)	0.009 (0.009)	0.003 (0.001)	0.001
PRESENT	FAR	LARGE	HIGH	0.052 (0.013)	0.003 (0.003)	0.004	0.004 (0.010)	0.010 (0.009)	-0.002 (0.002)	0.002
PRESENT	FAR	LARGE	HIGH	0.031 (0.006)	0.000 (0.001)	-0.000	-0.022 (0.005)	-0.307 (0.004)	0.000 (0.001)	-0.001

FIG 12   PAYOFF MATRIX FOR k=128, N=240 SHOWING ESTIMATED EXPECTED GAINS AND ASSOCIATED STANDARD ERRORS IN PARENTHESES

In this situation, the user could determine from Figure 7 that the search technique having the largest estimated expected gain (.910) was RSM-Variation I. Thus, a rational selection would be this search technique.

If, however, the true state of Nature is unknown, the user is faced with the quandry of how to exploit the data in the payoff matrix to make a good selection of a search technique. One possible solution is for him to adopt a maximin criterion[1] whereby he would select a search technique in such a way as *to* maximize the minimum gain which could be obtained. In general, this criterion results in what is known as a mixed strategy.[2] A mixed strategy defines a probability distribution over the set of search techniques, and the actual search technique to be used is based on a random selection governed by the probability distribution.

Application of the maximin criterion provides an overly conservative selection of search technique, however. Essentially, this criterion is based on the pessimistic assumption that Nature is a malicious opponent who will select a state in a conscious attempt to keep the simulation user's gain as low as possible. Since this pessimistic assumption is usually unwarranted, only in *extremely* rare instances would the user base his selection of a search technique on the maximin criterion.

Although the true state of Nature is usually, if not always, unknown to the user, he is generally not in complete ignorance because he tends to have some information about the simulation itself. Thus, the user is facing a condition of partial, rather than total, ignorance about the true state of Nature. Often he can summarize this partial ignorance by means of a probability distribution over the possible states of Nature. The existing information on which this probability distribution is based is often referred to as prior knowledge, and the resulting probabilities are known as prior probabilities. In those cases where no relative frequency basis exists for defining a probability distribution, "subjective" probabilities or "degrees of belief" may be used. Although much

---

[1] Because the entries in most payoff matrices are in terms of losses rather than gains, it is more or less standard to think of a minimax criterion. In the present situation where gains are being considered, a maximin criterion, the mirror-image of a minimax criterion, is appropriate.

[2] See, for example, Luce and Raiffa (1957).

controversy[1] exists as to the reasonableness and efficacy of subjective probabilities, it will be tacitly assumed that where no relative frequency based probabilities are available, the simulation user can specify the required probabilities on a subjective basis.

It is reasonable for the user to select that search technique which provides the maximum expected gain based on the prior probabilities[2]. Let the ith state of Nature be denoted by S_i, let the prior probability that S_i is the true state of Nature be denoted by p_i and let the estimated expected gain corresponding to the jth search technique[3] when S_i is the true state of Nature be denoted by G_{ij}. Then the search technique to be selected would be the technique which provided the largest value of $\sum_i p_i G_{ij}$. As a simple example, suppose that in a simulation situation with $k = 30$ and $N = 45$ the assigned probabilities were $p_5 = p_6 = p_7 = p_8 = .25$, with the remaining probabilities equal to zero. For these prior probabilities it can be easily verified from the payoff matrix in Figure 7 that the following estimated expected gains result:

.858 for Random Search

.390 for Single-Factor

.381 for Single-Factor accelerated

.920 for RSM-Variation I

.919 for RSM-Variation I accelerated

.880 for RSM-Variation II

.877 for RSM-Variation II accelerated.

Thus, the choice of a search technique would be RSM-Variation I, or perhaps its accelerated counterpart.

---

[1] See Good (1965) for a summary of various opinions in this controversy.

[2] This is known as a Bayes criterion.

[3] For convenience, the numbers of the states of Nature and those of the search techniques will be assumed to correspond to the order in which they are listed in the payoff matrices of Figures 5-12.

D. DISCUSSION OF RESULTS

The accumulated performance data provides a guide for selection of a search technique by the user of the prototype "OPTIMIZER" and for eventual incorporation of decision logic directly into the "OPTIMIZER". This guide is, admittedly, a rough one because of the two following reasons:

(1) Only two values of k were considered.

(2) Prior probabilities relating to the simulation characteristics are required.
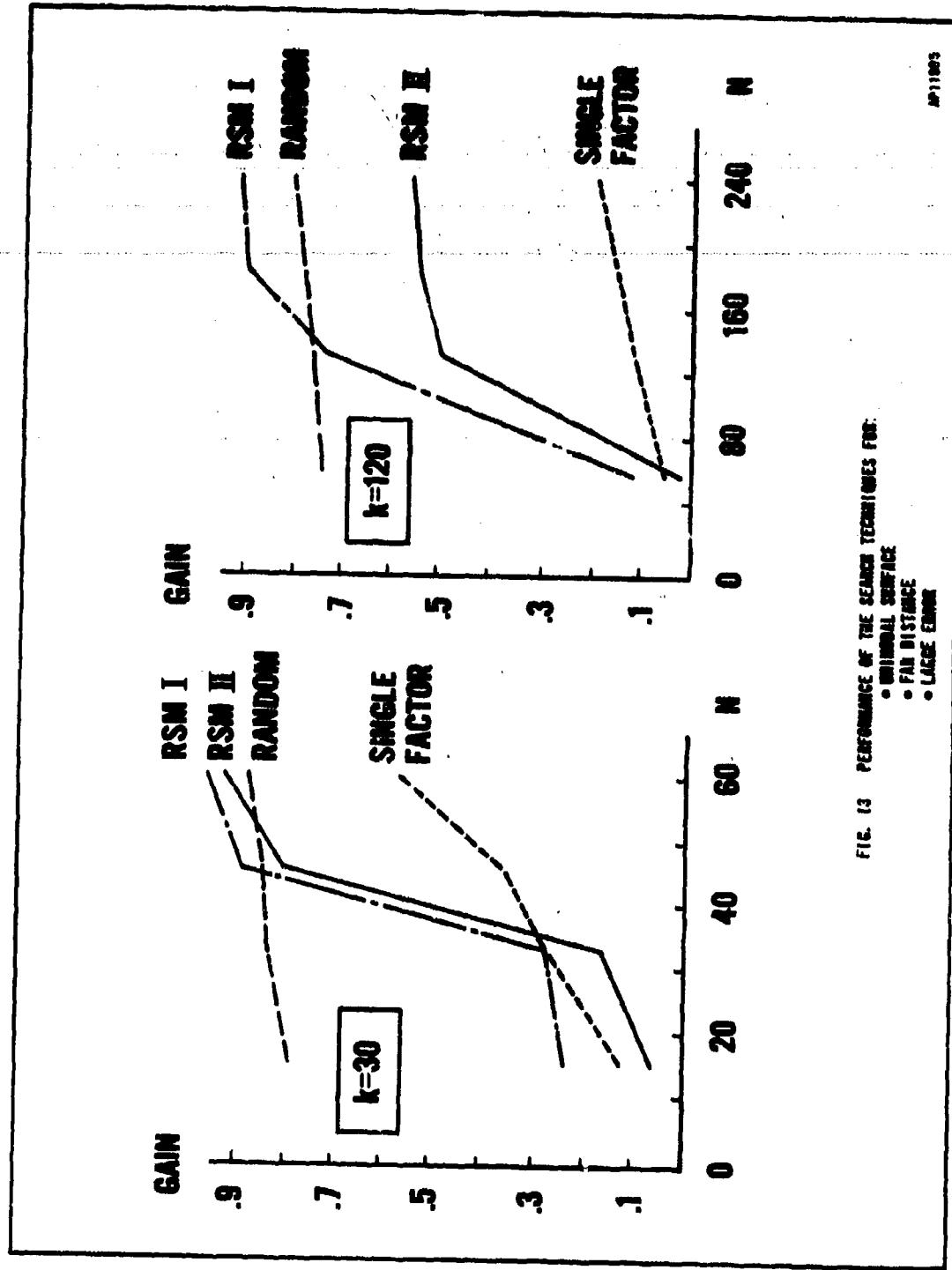
On this latter point, it should be noted that a good practice to follow in assigning subjective probabilities is to estimate upper and lower values for the probabilities[1]. If this is done, these values can be used in conjunction with the appropriate payoff matrix to determine the sensitivity of the search technique selection to the variation in the probabilities.

With regard to the former point, the results obtained for k = 30 and k = 120 may, in many instances, be cautiously interpolated and extrapolated to other values of k by comparing performance of the search techniques graphically. As a simple illustration, assume a simulation user who has a maximum of 90 computer runs available is certain that his simulation, which involves 60 controllable factors, can be categorized by a unimodal response surface (i.e., no local maxima), by a starting point far from the true optimum point, and by large random error. Furthermore, assume that he regards low activity or high activity as equally likely. In terms of the prior probabilities assigned to the states of Nature, his assignment would be $p_7 = p_8 = .5$, with all the other p_i's equal to zero. A plot of gain for each of the two values of k is given in Figure 13.

From this figure, it can be seen that the only two contending search techniques given the assigned prior probabilities are RSM-Variation I and Random Search.[2] Also, the user might reason that for both k = 30 and k = 120, the

---

[1] This suggestion is given by Good (1965), a firm believer in the usefulness of subjective probabilities.

[2] The accelerated techniques are not plotted because they provide essentially the same results as their nonaccelerated counterparts.

FIG. 13 PERFORMANCE OF THE SEARCH TECHNIQUES FOR:
• MINIMAL SURFACE
• FAR DISTANCE
• LARGE ERROR

best search technique when $N = 1.5 k$ is RSM-Variation I, and thus for $k = 60$ and $N = 90 = 1.5 k$, the best search technique would also be RSM-Variation I. It should be noted, however, that for k equal to 30, 60, or 120, the corresponding fractional factorial would require only a few runs more than k. If, for instance, $k = 64$, the corresponding fractional factorial would require twice this number of runs. This difficulty can be bypassed, at least for $k < 100$, by including the designs given by Plackett and Burman (1946) in the "OPTIMIZER". These designs, like the fractional factorials, involve only the values ± 1 in terms of the coded factors. Unlike the fractional factorials, the Plackett-Burman designs never require more than $k + 4$ computer runs. Although these designs[1] are not, in general, straightforward to construct, the relatively good performance of RSM-Variation I implies that they should definitely be incorporated into a final production version of the "OPTIMIZER".

It can be seen from the payoff matrices in Figures 5 through 12 that the performance of both Single-Factor techniques is quite close to being dominated by the performance of other techniques. Thus, for all practical purposes, these two techniques may be eliminated from consideration. Each of the remaining techniques, however, has the potential of providing the maximum gain, depending on the assignment of prior probabilities. Thus, these techniques should remain as options within "OPTIMIZER". It should be noted, however, the gains obtained for Random Search in the empirical study are likely to be larger than in practical situations because of the preferential definition of the overall search region discussed in Section III.C.1.

As an example, if each coordinate of the optimum point were in the interval $(-.75, .75)$, the overall search region would be given by $-1 \leq x_i \leq 1$ for $i = 1$, ..., k. Thus, the volume enclosed by this region would be 2^k. If the overall region were increased so that it were given by $-1.1 \leq x_i \leq 1.1$ for $i = 1, \ldots, k$, the volume enclosed would be $(2.2)^k$. The difference in the two volumes is extremely large for the values of k considered in the study. For instance, if $k = 120$, $2^k \doteq 1.33 \times 10^{36}$ while $(2.2)^k \doteq 1.23 \times 10^{41}$. As a result of increasing the interval on each x_i by 10% the corresponding volume of the overall search

---

[1] In actuality, the fractional factorial used for the case $k = 30$ is one of the family of Plackett-Burman designs.

region would increase by approximately 10,000,000% Thus, as the uncertainty in the interval associated with each factor increases slightly, the corresponding overall search region increases phenomenally, causing a decrease in the efficiency of Random Search. It is suggested, therefore, that the overall search region be carefully defined before applying Random Search. In any event, it can be seen from the data that Random Search should not necessarily be the search technique selected, even for a large number of factors. Thus, the current study has provided conclusions different from those of Brooks (1959) and McArthur (1961).

As far as the accelerated techniques are concerned, they do not appear to provide extremely different results than their nonaccelerated counterparts. In fact, paired-observation t-tests on the differences in the gains provided by each pair of accelerated and nonaccelerated techniques indicated no significant differences in performance.[1] However, from a decision-theoretic rather than a hypothesis testing viewpoint, it makes sense to consider both type of techniques, since neither dominates the other.

---

[1]Hindsight analysis indicates that while acceleration sometimes resulted in reaching a better response in a fewer number of computer runs, there were also many instances in which acceleration overshot, and wasted computer runs.

IV. APPLICATION OF THE PROTOTYPE "OPTIMIZER"
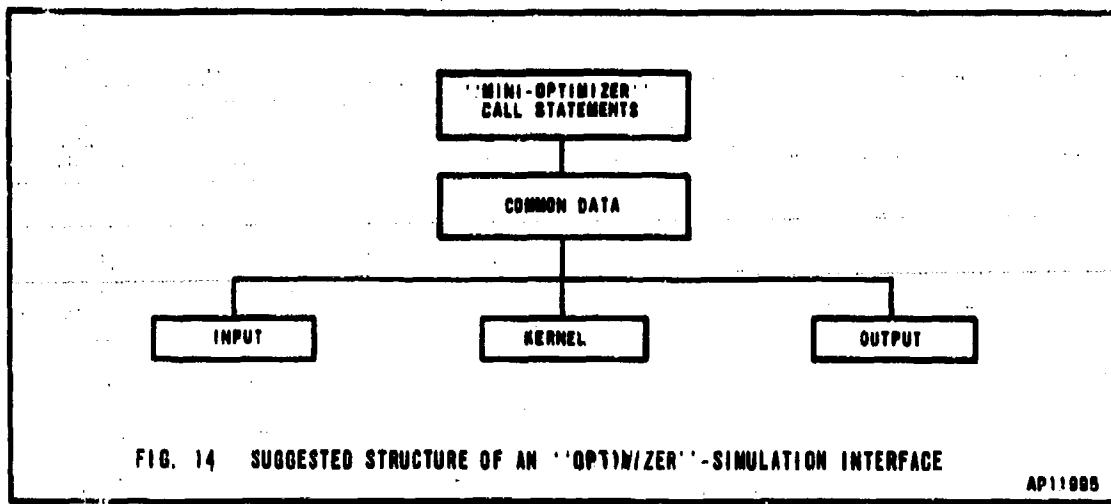
The potential interface with "OPTIMIZER" should, if possible, be considered during the development of a simulation. Every effort should be made to minimize the execution time of each simulation iteration, by eliminating unessential intermediate data analysis and summaries, for example. The input and output operations should be isolated, leaving only a complete portion of the model, or a "kernel." Making the kernel modular is highly desirable, expecially if overlay is necessary because of the size of the simulation.

Ideally, then, there would be three sections to every "OPTIMIZER" - simulation interface. Section I would contain the simulation input statements, logic checking routines, file and table creations, as well as the "OPTIMIZER" input statements. Operations in this section would be independent of operations in any other sections, although data would be exchanged with the other sections. Section II would contain the kernel and the "OPTIMIZER" search techniques. (*In* order to decrease core requirements, a "MINI-OPTIMIZER", i.e., one selected search technique, could be substituted.) Section III would contain the file update, report generator portions, "OPTIMIZER" summary output, etc., and would only be executed after Section II is completed.

Military simulations generally require large core storage requirements and execution time, and are usually Input/Output bound. It is not uncommon for one run to require 30 minutes to totally execute while 20 kernel runs can be executed in the same amount of time. Overlay structure is also quite common, giving added impetus to the three section structure as shown in Figure 14.

However, many candidate simulations for optimization are not under development now, but already exist. Hence, the interfacing may not be quite as simple, although the "MINI-OPTIMIZER"/kernel sectioning concept may still be a good one. An alternative, of course, is to relegate "OPTIMIZER" to a subroutine role, which would require only minor modification of "OPTIMIZER."

In order to gain more insight into the "OPTIMIZER"-simulation interface, as well as to examine the effectiveness of *the* prototype "OPTIMIZER", it seemed desirable to apply the "OPTIMIZER" in a realistic simulation situation. With the cooperation of the U.S. Army Strategy and Tactics Analysis Group (STAG), the "OPTIMIZER" was applied to a full-scale simulation. This test case application is described in the following section.

FIG. 14 SUGGESTED STRUCTURE OF AN "OPTIMIZER"-SIMULATION INTERFACE

AP11985

A. A TEST CASE: FORECAST II

FORECAST II is a one-sided expected value model which attempts to tactically allocate finite resources against an enemy force. These resources are airborne delivery and weapon systems combinations which are offensive, offensive-defensive or defensive in nature. The enemy force is represented by a strike list. Given a strike list, the model attempts to achieve a specified assurance of damage while expending a minimum number of offensive sorties, thus allowing more targets to be engaged. The descriptors within the model are specifically related to the military capabilities of the delivery and weapon systems. The input variables which may be considered controllable by the military decision makers are those describing the distribution of the delivery and weapon systems of the friendly force and those representing the friendly force's strategy in the form of an ordered strike list.

The prototype "OPTIMIZER" was employed to study the distribution of the resources for a fixed strike list. Thus, the results of the study were unique to the strike list in effect. The figure of merit, selected to conform to the model action, was chosen as "the number of targets attacked per sortie dispatched." The factors selected for optimization were those parameters describing the distribution of aircraft to base zones. Five aircraft types were simulated, of which four types were permitted in three base zones, while the other was allowed in only two zones. The resulting 14 controllable factors were subject to constraints imposed by a constant total inventory for each type of aircraft.

-320-

1. Interfacing

FORECAST II is a computer simulation, written in FORTRAN, which
fits the general description previously attributed to military models; i.e., it
requires large amounts of core storage and generates large amounts of
ancillary information. Preliminary discussion with the analyst who developed
FORECAST II (a distinct advantage when attempting to interface "OPTIMIZER"
with an existing simulation) led to the conclusion that a kernel model coupled
with a "MINI-OPTIMIZER" would provide the best mode of interfacing. It was
jointly decided that:

(a) Unessential operations such as collateral damage assessment,
normal output operations, etc., should be deleted in the kernel to
reduce machine time during execution.

(b) The link structure should be modified to make additional core
storage available and the "OPTIMIZER" should be modified to
delete unapplicable options and reduce core storage requirements.

(c) A fully edited binary tape containing the strike list should be
created in place of a BCD tape in order to decrease processing
time.

These modifications and several other refinements were made to the
"OPTIMIZER" and kernel in order to bypass resetting inventories, to calculate
the figure of merit directly, and to speed processing. The resulting computer
programs, which provided the required interface, were successfully executed
and used throughout the study.

2. Two Problems

The controllable factors (the resource distribution parameters) violated
the basic assumptions on which the prototype "OPTIMIZER" was constructed.
These assumptions are that the controllable factors are continuous and un-
constrained. Unfortunately, the resource distribution parameters were non-
negative integer values which had to total to a set integer value. It was felt that
the violation of the assumption of continuity could be ignored with little impact
on the "OPTIMIZER" results since the magnitude of the total inventory was
sufficiently large to make transition between integers small.

-321-

However, the prototype "OPTIMIZER" is not capable of handling constraints, while the controllable factors required two types:

(a) $X_i \geq 0$

(b) $\Sigma\ X_i = C$

Hence, it was decided to develop a reparameterization in an attempt to overcome the constraint difficulty. The specific reparameterization selected is illustrated in the following paragraph for the case where four factors are involved.

Consider four continuous factors -- X_1, X_2, X_3, X_4 -- subject to the above constraints. It should be noted that if the values of three of these factors are known, then the fourth is automatically determined by constraint (b). Now consider the following reparameterization:

$$W_1 = \ln\ (X_1/X_4)$$
$$W_2 = \ln\ (X_2/X_4)$$
$$W_3 = \ln\ (X_3/X_4)$$

The reparameterized factors W_1, W_2, and W_3 are continuous and unconstrained. The inverse of this reparameterization, together with constraint (b), results in:

$$X_1 = Ce^{W_1}/(1 + \sum_1^3 e^{W_i})$$

$$X_2 = Ce^{W_2}/(1 + \sum_1^3 e^{W_i})$$

$$X_3 = Ce^{W_3}/(1 + \sum_1^3 e^{W_i})$$

$$X_4 = C - (X_1 + X_2 + X_3)$$

It is thus possible to encode the constrained X factors, optimize in the W factor space, and decode the "optimum" point to obtain the values of $(X_1,\ X_2,\ X_3,\ X_4)$.

This type of reparameterization reduced the set of 14 original factors to a set of only 9 new factors. Although dealing with the factors arising from

the transcendental reparameterization posed no problems to the "OPTIMIZER",
the simulation user experienced some difficulty in the interpretation of the
optimization of the transformed factors. Specifically, he used a few runs to
apply the Single-Fac · search technique in order to check whether the kernel
model, the "MINI-C TIMIZER", and the interfacing were working properly, and
whether the constraints were being observed. The difficulty arose because the
change of a single W factor causes simultaneous changes in the X factors.
Because a simulation user would, in general, feel more comfortable with
optimization in terms of factors which are directly related to quantities in the
real world, reparameterization must be regarded as a stopgap measure, at best.

B. DISCUSSION OF RESULTS

After the interfacing was checked and the reparameterization programmed
into the "MINI-OPTIMIZER"/kernel, the analyst provided a starting point which
he regarded as a "pretty good" allocation. After the step sizes and the overall
search region were selected, four searches were made. These searches were
based on Random Search, Single-Factor, RSM-Variation I, and RSM-Variation
II, respectively. Each search involved k=9 transformed factors and was per-
mitted a maximum of N=20 computer runs.

1. Performance of the Search Techniques

The specified starting point provided an observed response of .397.
Because the simulation is based on an expected value model, this observed
response is also the true response at that point. Thus, in terms of the measure
used in the empirical studies, the relative gain provided by a search technique
was given by:

$$\left\{ \frac{(\text{"Optimum" response found by technique}) \ - \ .397}{(\text{True optimum response}) \ \qquad \ - \ .397} \right\}$$

Because the true optimum is unknown, the best that could be done was to
determine a lower bound for relative gain. This lower bound was calculated
using information from the analyst who verified on theoretical grounds that the
figure of merit obtained in the simulation would never be greater than .500.
This value was used as the true optimum response in the definition of relative
gain, yielding the lower bound.

The results of the application of the four search techniques to FORECAST II are given in *Figure* 15. As can be noted, RSM-Variation I provided the best results, a relative gain of greater than 38% in just 20 computer runs. Somewhat surprising was the performance of the Single-Factor approach and RSM-Variation II. On the basis of results in the empirical studies, it might have been anticipated that Single-Factor would not have performed as well as it did. Likewise, it might have been anticipated that the performance of RSM-Variation II would have been better. It must be borne in mind, however, that the results of the empirical studies provided the value of estimated expected gain based on a number of iterations of the same search technique. In the FORECAST II application, the results are based on only a single iteration of each search technique.

As a whole, the application of the prototype "OPTIMIZER" to FORECAST II proved a success. The relative ease of interfacing and using the "OPTIMIZER", combined with the increased figure of merit found by the "OPTIMIZER", shows the potential value of such an externally-controlled optimization computer program. However, it became apparent during the "OPTIMIZER" application to FORECAST II that two modifications to "OPTIMIZER" were extremely desirable.

2. Two Indicated Modifications to "OPTIMIZER"

Experience with application of the prototype "OPTIMIZER" to FORECAST II revealed that a modification to permit constrained optimization problems and one to provide a restart capability would provide a more powerful "OPTIMIZER."

To deal with constraints using the prototype "OPTIMIZER", it may be possible to change the problem into an unconstrained form either by means of an appropriate reparameterization on the controllable factors or by modifying the simulation to provide a "bad" figure of merit value for any combination of controllable factors which violates the constraints.[1] However, neither of these procedures guarantee a solution to the constraint problem and, even if successful, are at best stopgap measures which levy additional requirements on the user of

---

[1] Both of these stratagems appeared to be reasonably acceptable in the application of the "OPTIMIZER" to FORECAST II.

TECHNIQUE	RESPONSE AT STARTING POINT	"OPTIMUM" RESPONSE FOUND	LOWER BOUND FOR RELATIVE GAIN
RANDOM	.397	.423	25%
SINGLE-FACTOR	.397	.427	29%
RSM I	.397	.436	38%
RSM II	.397	.417	19%

FIG. 15   RESULTS OF APPLICATION OF PROTOTYPE "OPTIMIZER"

AP11986

the "OPTIMIZER." Thus, "OPTIMIZER" should be modified to allow a reasonable number of constraints to be handled without any intervention of the user, except for specification of the constraints as data input.

The "OPTIMIZER" logic was designed to make the most efficient use of the available number of computer runs. This logic was based on the hypothesis that the user would assign to the "OPTIMIZER" the maximum number of computer runs that he could justify. For example, assume that one of the RSM techniques is used and a point is reached where a new experimental design is required to determine a new path of steepest ascent. If not enough computer runs remain to permit a full design involving the original k controllable factors, the "OPTIMIZER" uses a full design with only a selected subset of the k factors, instead of beginning a portion of the design involving all of the controllable factors. This procedure permits a path of steepest ascent (in a space of reduced dimensionality) to be calculated and followed.

Based on the "OPTIMIZER" application to FORECAST II, it is apparent that in many situations the user may not be able to (or may not wish to) use the total number of available runs in one "OPTIMIZER" run. Thus, the user might desire to restart the "OPTIMIZER" from the point which its original application found to be "optimum" within the number of computer runs originally specified. The current version of the "OPTIMIZER" does not permit such a continuation.

In order to provide efficient restart capabilities, the "OPTIMIZER" should be modified so that it "remembers" the phase of the "optimization" process in progress when the specified number of computer runs was exhausted. This modification would assure that all pertinent data (such as maximum response observed, location of that response, parameters determining the path of steepest ascent) are available should the user desire to continue the "optimization" process from where it had originally terminated.

V. CONCLUSIONS

In summary, the results of the HRB-Singer study to date have been three-fold:

    (1) A prototype "OPTIMIZER" exists.

    (2) Information to aid in the selection of a search technique is available.

    (3) Preliminary application of the prototype "OPTIMIZER" has given positive results.

At this stage of the study, it seems plausible that the "OPTIMIZER" may prove of great value to the simulation user and, hence, its development should be continued.

Despite a growing literature[1] which touches on the topic of optimization in the simulation framework, most practical applications tend to rely on trial-and-error procedures or on the relatively inefficient Single-Factor approach. Thus, the primary utility of the "OPTIMIZER" will be that it provides the simulation user with a means of managing an efficient automatic search for an optimum.

In the rare case where familiar processes are being simulated, where a well-understood figure of merit is to be optimized, and where only a few factors affect this figure of merit, the simulation user may be able to do a fairly good job of optimizing by heuristic means. In essence, he would select initial values for the controllable factors, run the simulation for these values, observe intermediate output and the figure of merit obtained, use this data to form tentative hypotheses about how the values of the controllable factors should be changed, change these values, and keep repeating this procedure. Needless to say, much effort would be required of the user, and the whole process could easily extend over weeks because of the time required for analysis and for computer turnaround between the individual runs.

---

[1] For example, Ackoff (1962), Bonini (1963), Burdick and Naylor (1966), Conway, Johnson, and Maxwell (1959), Emshoff and Sisson (1970), Hunter and Naylor (1970), Jacoby and Harrison (1962), Mihram (1970), Naylor (1969), Naylor, Balintfy, Burdick, and Chu (1966), Overholt (1970), Smith (1969), Smith (1970a), and Smith (1970b).

Application of the "OPTIMIZER", on the other hand, would require little effort on the part of the user, except for the initial interfacing. Further, by automating intermediate analyses, and thereby consolidating the many runs otherwise necessary, the "OPTIMIZER" would significantly reduce the overall time involved. Therefore, even in this situation, the "OPTIMIZER" has potential usefulness.

As an alternative to an "OPTIMIZER" computer program, a set of search technique alogrithms could be provided to the simulation user. However, the fairly tedious steps in some of these algorithms, coupled with the fact that simulation input would have to be manually changed between each computer run, might prove discouraging to the user, causing him to abandon the algorithms in favor of an easier, but less efficient approach. If, in fact, the user did persevere with the algorithms, it is reasonable to assume that he would write his own program to carry out the search technique steps and to permit automatic modification of simulation inputs. Thus, supplying the simulation user with only the algorithms appears highly inefficient compared to providing him with a proven "OPTIMIZER" computer package.

It must be noted that further work remains to be done before the "OPTIMIZER" can be used as an effective tool in practical applications. For example, the prototype should be modified to provide the restart and constrained optimization capabilities discussed in the previous chapter. In addition, further investigative applications of the "OPTIMIZER" to full-scale computer simulations should be made in order to provide information for a detailed critique of its overall performance.

Nonetheless, there are strong indications at the present time that the "OPTIMIZER" should prove valuable to the simulation user.

VI. BIBLIOGRAPHY

Ackoff, R. L. (1962). Scientific Methods: Optimizing Applied Research Decisions, John Wiley and Sons, New York.

Bonini, C. P. (1963). Simulation of Information and Decision Systems in the Firm, Prentice-Hall, Inc., Englewood Cliffs, N. J.

Box, G. E. P. (1952). "Multi-factor Designs of First Order," Biometrika, Vol. 39, p. 49.

Box, G. E. P., and Wilson, K. B. (1951). "On the Experimental Attainment of Optimum Conditions," Journal of the Royal Statistical Society, (Series B), Vol. 13, p. 1.

Brooks, S. H. (1959). "A Comparison of Maximum - Seeking Methods," Operations Research, Vol. 7, p. 430.

Burdick, D. S. and Naylor, T. H. (1966). "Design of Computer Simulation Experiments for Industrial Systems," Communications of the Association for Computing Machinery, Vol. 9, No. 5, p. 329.

Cochran, W. G., and Cox, G. M. (1957). Experimental Designs, (2nd Edition), John Wiley and Sons, Inc., New York.

Conway, R. W., Johnson, B. M., and Maxwell, W. L. (1959). "Some Problems of Digital Systems Simulation," Management Science, Vol. 6, p. 92.

Davies, O. L., ed. (1967). Design and Analysis of Industrial Experiments, Halfner Publishing Co., New York.

Emshoff, J. R. and Sisson, R. L. (1970). Design and Use of Computer Simulation Models, The MacMillan Company, New York.

Friedman, M., and Savage, L. J. (1947). Chapter 13, Techniques of Statistical Analysis, Eisenhart, C., Hastay, M., and Wallis, W., McGraw-Hill Book Co., New York.

Good, I. J. (1965). The Estimation of Probabilities: An Essay of Modern Bayesian Methods, M. I. T. Press, Cambridge, Mass.

Hillier, F. S., and Lieberman, G. J. (1967). Introduction to Operations Research, Holden-Day, Inc., San Francisco.

Hooke, R., and Jeeves, T. A. (1958). "Comments on Brooks' Discussion of Random Methods," Operations Research, Vol. 6, p. 881.

Hunter, J. S., and Naylor, T. H. (1970). "Experimental Designs for Computer Simulation Experiments," Management Science, Vol. 16, p. 422.

VI. BIBLIOGRAPHY (Cont'd)

Jacoby, J. E., and Harrison, S. (1962). "Multi-Variable Experimentation and Simulation Models," Naval Research Logistics Quarterly, Vol. 9, p. 121.

Kendall, M. G., and Stuart, A. (1963). The Advanced Theory of Statistics, Vol. I., Hafner Publishing Co., New York.

Luce, R. D., and Raiffa, H. (1957). Games and Decisions, John Wiley and Sons, New York.

McArthur, D. S. (1961). "Increasing the Efficiency of Empirical Research," Esso Research and Engineering Company, Report No. RL34M61.

Mihram, G. A. (1970). "An Efficient Procedure for Locating the Optimal Simular Response," Proceedings, Fourth Conference on Applications of Simulation, New York.

Naylor, T. H., ed. (1969). The Design of Computer Simulation Experiments, Duke University Press, Durham, N. C.

Naylor, T. H., Balintfy, J. L., Burdick, D. S., and Chu, K. (1966). Computer Simulation Techniques, John Wiley and Sons, Inc., New York.

Overholt, J. L. (1970). "Analysis, Data Inputs and Sensitivity Tests in War Games," Professional Paper No. 13, Center for Naval Analysis. (AD722-858).

Plackett, R. L., and Burman, J. P. (1946). "The Design of Optimum Multi-factor Experiments," Biometrika, Vol. 33, p. 305.

Scheffé, H. (1959). The Analysis of Variance, John Wiley and Sons, New York.

Smith, D. E. (1969). "Optimization Within Computer Simulation: An Application of Response Surface Methodology," 24th Military Operations Research Symposium Proceedings, p. 181.

Smith, D. E. (1970a). "An 'OPTIMIZER' for Use in Computer Simulation: Background and Design Concepts," HRB-Singer Report No. 4352.11-R-1. (AD708-765).

Smith, D. E. (1970b). "Optimization in Computer Simulation: Some General Comments," 38th National Meeting of the Operations Research Society of America. (Abstract in ORSA Bulletin, Vol. 18, No. 2., p. 163).

Winer, B. J. (1962). Statistical Principles in Experimental Design, McGraw-Hill Book Co., New York.

STAG MONOTONE EXPERIMENTAL DESIGN ALGORITHM (SMEDAL)

PFC Alexander Morgan
Systems Development Division
US ARMY STRATEGY AND TACTICS ANALYSIS GROUP
Bethesda, Maryland

ABSTRACT

Experimental design techniques are used throughout the sciences in the analysis of statistical processes. These techniques can also be used to analyse large, essentially deterministic, complex systems. In this non-statistical case, they are not necessarily the most efficient tools.

In the present paper we describe a method which was developed to aid in the investigation of a large deterministic computer simulation. This method would seem to be applicable to the study of many complex non-statistical systems. We include possible extensions of the method as well as several alternate approaches to the general problem. Further work to extend this technique to certain statistical systems is planned.

INTRODUCTION

Experimental design involves the analysis of a complex system by choosing a series of experiments to perform on the system. Techniques for analysis of statistical processes have been developed including factorial, fractional factorial, and other such designs. (Reference [1] by Hunter and Naylor gives a good summary of these.) For systems with more than a few inputs or for which experiments tend to be expensive a random search often gives the most efficient design, and its efficiency can be quite low. (See McArthur [3].)

In this paper we describe a technique (called the STAG Monotone Experimental Design Algorithm (SMEDAL)) for analysing deterministic, rather than statistical, systems which exhibit certain properties (described in section I). SMEDAL was originally developed to aid in the analysis of computer simulations and seems to be applicable to many systems in which statistical variation plays a minor, or non-existent role. Extension of this technique to statistical systems is planned.

The remainder of this paper has been photographically reproduced from the manuscript submitted by the author.

SECTION I
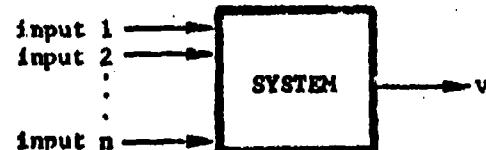
GENERAL PROBLEM

We have a complex system with discrete inputs and outputs. We wish to analyse the relations between inputs and outputs with the general purpose being to determine the response of the system using the smallest number of actual experiments possible. Specific relations between inputs are considered and their usefulness to this program evaluated.

SPECIFIC PROBLEM

The system takes n inputs where input i has m_i levels. Its output is v where v = 0,1,2, ... k. (See figure 1.)

Figure 1

$$1 \leq input\ i \leq m_i \qquad 0 \leq v \leq k$$

A set of inputs for one experiment will be considered to be an n-tuple $(x_1, \ldots ,x_n)$ with $1 \leq x_i \leq m_i$ for all i. We will say that $N(x_1, \ldots ,x_n) = v$ if experiment $(x_1, \ldots ,x_n)$ yields output v. Therefore N is a function from the $m_1 m_2 \ldots m_n$ n-tuples to the set $\{0,1,2, \ldots ,k\}$.

The first relation between inputs is the <u>monotone relation</u>.

This relation states that if all inputs but one are held fixed, then an increase in the unfixed input cannot cause a decrease in the output. In symbols:

If $N(x_1, \ldots, x_n) = v$, then $N(x_1, \ldots, x_i+1, \ldots x_n) \geq v$ for all i and $N(x_1, \ldots, x_i-1, \ldots, x_n) \leq v$ for all i.

Example: Let $n = 8$ and $m_i = 3$ for all i. Let $k = 1$. If $N(2,2,\ldots,2) = 1$, then $N(3,2,2,\ldots,2) = 1$. In fact, if $(x_1, \ldots, x_n)$ is any combination of 2's and 3's, then $N(x_1, \ldots, x_n) = 1$. If $N(2,2,\ldots,2) = 0$, similar results follow as above with "1" substituted for "3". Thus it follows that $N(1,2,2,\ldots,2) = 0$, etc.

In the above example, if $N(2,2,\ldots,2) = 0$, then the monotone relation tells nothing about $N(1,3,2,\ldots,2)$. One input has been raised and another lowered; monotonicity does not apply in such a case. Although much can be done with our problem using only the monotone relation, it is possible that there will also exist dominance relations between the positions of the inputs. These dominance relations will occur when it is uniformly true that raising certain inputs and lowering certain others always tends to have the same effect on the output.

First, to clarify what is meant by "position", consider the 8-tuple $(3,2,1,1,1,1,1,1)$. 3 is in the first position, 2 is in the second position, and 1's are in positions 3 through 8.

Say that position i dominates position j if an increase of 1 in position i with a decrease of 1 in position j cannot cause a decrease in the output. In symbols:

Position i dominates position j means if $N(x_1, \ldots, x_n) = v$, then $N(x_1, \ldots, x_i+1, \ldots, x_j-1, \ldots, x_n) \geq v$ and $N(x_1, \ldots, x_i-1, \ldots, x_j+1, \ldots, x_n) \leq v$.

This relation could occur between any two positions. The more such relations that exist, the more information that can be obtained.

Example: If position 1 dominates position 2 and $N(1,3,3,2,2,1,1,1) = 1$, then $N(2,2,3,2,2,1,1,1) = 1$ (using parameters of previous example).

SECTION II

OUTLINE OF THE ALGORITHM

A detailed description of SMEDAL, with possible extensions, is given in the following sections. Included here is a basic outline of the algorithm and its use.

Consider two computer programs called, for the purposes of this paper, P and Q.

1. P has as inputs: (1) the dominance relations for the specific system under study (There possibly are none.), (2) a set S_i of n-tuples (representing those tuples c for which $N(c)$ is not known), and (3) tuple a_{i+1} with its outcome, $N(a_{i+1})$. (The tuple a_{i+1} represents the last experiment run on the system; $N(a_{i+1})$ has just been determined.) P has for output the set $S_{i+1} = S_i - K_{i+1}$ where K_{i+1} consists of all n-tuples $b_1, \ldots, b_r$ in S_i for which $N(a_{i+1})$ implies $N(b_1), \ldots, N(b_r)$, using

the monotone and dominance relations. Thus S_{i+1} contains everything in S_i that is not in K_{i+1} . (See figure 2.)

dominance relations ⟶

set S_i of tuples still undetermined ⟶ **P** ⟶ S_{i+1} , the new set of tuples still undetermined

tuple a_{i+1} with $N(a_{i+1})$
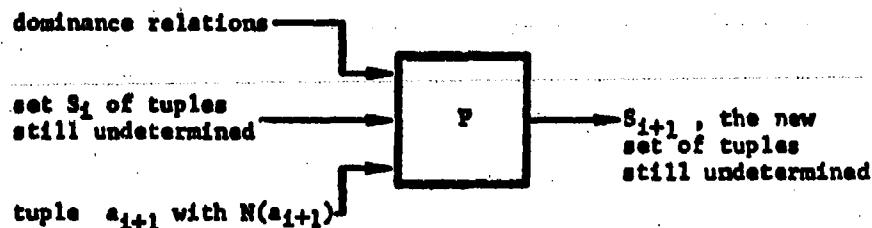
Figure 2

2. Q has as inputs the dominance relations (if any) and the set S_j of tuples still undetermined. Q's output is one tuple, a_{j+1} from S_j , where a_{j+1} is chosen so that $N(a_{j+1})$, whether $N(a_{j+1}) = 0$ or 1 or . . . or k, will determine the maximum number of tuples x in S_j . (See figure 3.)

dominance relations ⟶

set S_j of tuples still undetermined ⟶ **Q** ⟶ tuple a_{j+1} , the "best" tuple to run next
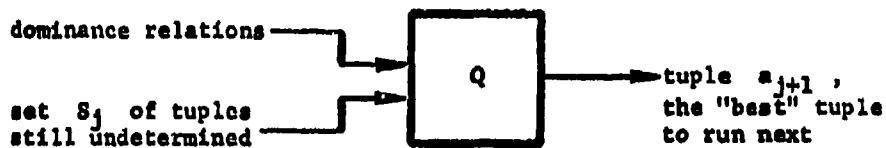
Figure 3

Using P and Q , the algorithm works as follows: (See figure 4, next page.)

(1) With S = all tuples, run Q to get a_1.

(2) Run a_1 . Find $N(a_1)$. (first experiment)

(3) With S = all tuples and $N(a_1)$, run P to get $S_1 = S - K_1$, where K_1 = set of tuples $b_1, \ldots, b_r$ for which $N(a_1)$ implies $N(b_1), \ldots, N(b_r)$.

(4) With S_1, run Q to get a_2.

(5) Run a_2. Find $N(a_2)$. (second experiment)

(6) With S_1 and $N(a_2)$, run P to get $S_2 = S_1 - K_2$, where K_2 is as above.

(7) With S_2, run Q to get a_3.

.

.

.

(8) Continue until all tuples are determined or further experiments are uneconomical.

$S \longrightarrow [Q] \longrightarrow a_1 \longrightarrow [SYSTEM]$

$\downarrow$

$N(a_1)$

$\downarrow$

$S \longrightarrow [P]$

$\downarrow$

$S_1 \longrightarrow [Q] \longrightarrow a_2 \longrightarrow [SYSTEM]$

$\downarrow$

$N(a_2)$

$\downarrow$

$S_1 \longrightarrow [P]$

$\downarrow$

$S_2 \longrightarrow [Q] \longrightarrow a_3 \longrightarrow [SYSTEM]$

.

.

.
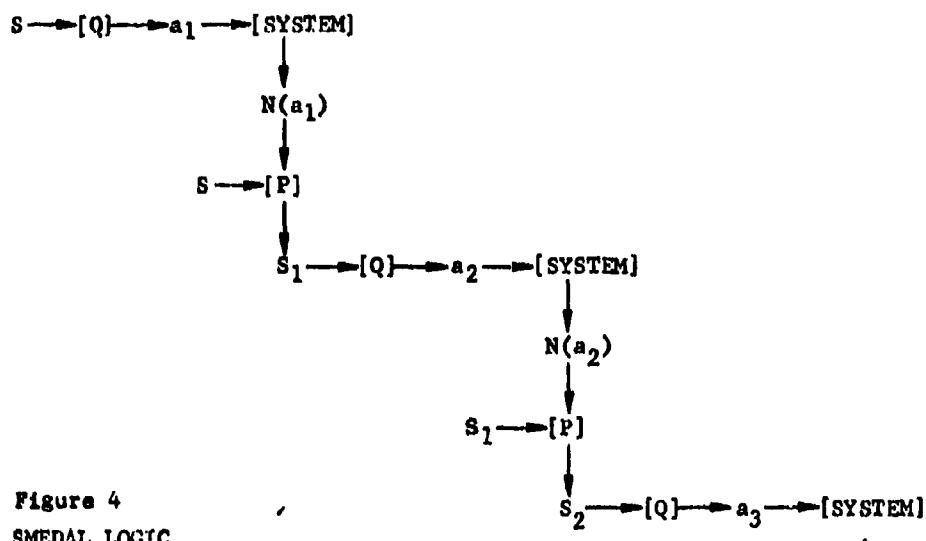
Figure 4
SMEDAL LOGIC
FLOW DIAGRAM

A number of special cases of SMEDAL have been run to help estimate the efficiency of the algorithm. In general, the number of system experiments was reduced to less than 20% of the total number to get all outcomes, and to a much smaller percent to get less than all outcomes.

The following table summarizes some of these runs. The case of 6 inputs at 3 levels each, with two outputs (0 and 1), was chosen as representative, although other cases have been run with comparable results.

TABLE

6 inputs at 3 levels each, 2 outputs

729 outcomes to determine (total)

For each case, the values $N(a_i)$ for $i = 1, 2, \ldots$ were chosen at random. $Q(w, z)$ with $w = 2$ and $z = .3$ was used. (See Section III, part D.)

Case Number	Number of System Experiments made	Percent of Number of System Experiments made Over total number	Percent of all Outcomes Determined
1	12	1.6%	50%
	19	2.6%	60%
	22	3.0%	70%
	33	4.5%	80%
	65	8.9%	90%
	140	19.2%	100%

Case Number	Number of System Experiments made	Percent of Number of System Experiments made Over total number	Percent of all Outcomes Determined
2	6	.8%	50%
	9	1.2%	60%
	14	1.9%	70%
	19	2.6%	80%
	45	6.2%	90%
	121	16.6%	100%
3	11	1.5%	50%
	15	2.1%	60%
	17	2.3%	70%
	27	3.7%	80%
	46	6.3%	90%
	102	14.0%	100%
4	11	1.5%	50%
	14	1.9%	60%
	18	2.5%	70%
	26	3.6%	80%
	42	5.8%	90%
	100	13.7%	100%

P AND Q FOR k = 1

This section contains the details of programs P and Q,
whose use is described in section II, for the case k = 1. In
this case, the output of the system is 0 or 1 .

PART A

First it will be helpful to make some definitions and
establish some propositions.

1. **Definition**. Let a be an n-tuple. Then there are collections
of n-tuples $\{b_1, \ldots, b_r\}$ and $\{c_1, \ldots, c_s\}$ such that
if $N(a) = 1$, then by the monotone relation $N(b_i) = 1$ for
$i = 1, 2, \ldots, r$, and if $N(a) = 0$, then by the monotone relation
$N(c_j) = 0$ for $j = 1, 2, \ldots, s$.   Call r+1 "the number of outcomes
determined if a wins" and denote it by $\#(a)$.   Call s+1 "the
number of outcomes determined if a loses" and denote it by
$\#'(a)$.

Example: If the system has 8 inputs at three levels each,
then $\#(22222222) = 256$ and $\#'(12222222) = 128$.   (Note: tuples
$(2,2,2,2,2,2,2,2)$ will be written (22222222) from now on.)

2. <u>Definition</u>.   Let $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$
be tuples. Then $lub(x,y) = z$ where $z = (z_1, \ldots, z_n)$ with
$z_i = max(x_i, y_i)$, and $glb(x,y) = w$ where $w = (w_1, \ldots, w_n)$
with $w_i = min(x_i, y_i)$.

Example: Let $x = (31222222)$ and $y = (22212222)$. Then $\text{lub}(x,y) = (32222222)$ and $\text{glb}(x,y) = (21212222)$.

3. **Proposition.** $\#(x \text{ and } y) = \#(x) + \#(y) - \#(\text{lub}(x,y))$ and $\#'(x \text{ and } y) = \#'(x) + \#'(y) - \#'(\text{glb}(x,y))$ where $\#(x \text{ and } y)$ is the number of outcomes determined if $N(x) = 1$ and $N(y) = 1$.

Proof: Immediate from definitions and monotone relation.

Example: Suppose $N(22222222) = 1$ and $N(31222222) = 1$. Since $\#(22222222) = 256$ and $\#(31222222) = 192$, one might guess (or hope) that knowing $N(22222222) = 1$ and $N(31222222) = 1$ would give $256 + 192 = 448$ outcomes. However, this is not true. Since $\text{lub}((22222222),(31222222)) = (32222222)$ and $\#(32222222) = 128$, proposition 3 implies we have only $256 + 192 - 128 = 320$ outcomes.

4. We will now describe program P. For simplicity, assume (for the moment) that there are no dominance relations. To start, P will print table T_0, which consists of rows of the form: x, $\#(x)$, $\#'(x)$ for all tuples x. In this case $\#(x)$ and $\#'(x)$ are easy to compute by a simple application of monotonicity.

Example: Assume 8 inputs at 3 levels each. Then $\#(22222222) = 2^8 = 256$, $\#(32222222) = 2^7 = 128$, $\#(12222222) = 3 \cdot 2^7 = 384$, $\#'(22222222) = 2^8 = 256$, $\#'(32222222) = 3 \cdot 2^7 = 384$, $\#'(12222222) = 2^7 = 128$, $\#(31222222) = 3 \cdot 2^6 = 192$, and $\#'(11322222) = 3 \cdot 2^5 = 96$.

After program Q has chosen a_1 and $N(a_1)$ has been determined, P prints table T_1, which consists of rows of the form: x, $\#_1(x)$, $\#'_1(x)$ where $\#_1(x)$ and $\#'_1(x)$ are computed as follows:

If $N(a_1) = 0$, then $\#_1(x) = \#(x)$ and $\#'_1(x) = \#'(x) - \#'(glb(a_1,x))$.

If $N(a_1) = 1$, then $\#_1(x) = \#(x) - \#(lub(a_1,x))$ and $\#'_1(x) = \#'(x)$.

In general, after Q has chosen a_i and $N(a_i)$ has been determined, P prints table T_i which consists of rows of the form:

x , $\#_i(x)$, $\#'_i(x)$ where: If $N(a_i) = 0$, then $\#_i(x) = \#_{i-1}(x)$ and $\#'_i(x) = \#'_{i-1}(x) - \#'_{i-1}(glb(a_i,x))$. If $N(a_i) = 1$, then $\#_i(x) = \#_{i-1}(x) - \#_{i-1}(lub(a_i,x))$ and $\#'_i(x) = \#'_{i-1}(x)$.

5.   It is easy to verify that the rows of table T_i give: tuple x, the number of outcomes determined if x wins, the number of outcomes determined if x loses, given that $a_1, \ldots , a_i$ have been chosen and $N(a_1), \ldots , N(a_i)$ have been already determined. (If $N(x)$ is determined by a_r, then either $\#_j(x)$ or $\#'_j(x)$ will be 0 for $j \geq r$. These tuples x will be automatically excluded from further consideration, if the method for computing the T_i described above is used. See Part C concerning program Q. Set S_i, referred to in section II, consists of those tuples x for which either $\#_j(x)$ or $\#'_j(x)$ equals 0.) Note that proposition 3 implies: if $N(a_1) = 1$, then $\#_1(x) = \#(x$ and $a_1) - \#(a_1) = \#(x) - \#(lub(a_1,x))$, and clearly this formula iterates with $\#_i$ and a_i replacing $\#_1$ and a_1.

It is possible to write down a non-iterative formula for $\#_i(x)$ using i_0 nested summations and the i_0^{th} generalization of proposition 3 where i_0 is the number of a_j such that $\#(a_j) = 1$ for $j < i$, but there seems to be no real use for this generalized formula, considering the sequential nature of SMEDAL.

PART B

Now we will consider how P is to be adjusted to take into account dominance relations. First, some definitions and their consequences will be stated.

6. __Definition__. Let x and y be tuples. We say "x *dominates* y" if $N(x) \geq N(y)$. In this case we write $x \gg y$. (Note: "$\gg$" is a partial order on the set of tuples.)

7. __Proposition__. $x \gg y$ if and only if $N(y) = 1$ implies $N(x) = 1$ __and__ $N(x) = 0$ implies $N(y) = 0$.

8. __Proposition__. $x \gg y$ if $x_i \geq y_i$ for all i.

9. __Proposition__. If position i dominates position j (terminology from section I), $x_r \geq y_r$ for all $r \neq i,j$, and $x_i - 1 \geq y_i$, $x_j + 1 \geq y_j$, then $x \gg y$.

Proof: We use proposition 7. Suppose $N(y) = 1$. Then

$$1 = N(y_1, \ldots, y_n) \leq N(y_1, \ldots, x_i - 1, \ldots, x_j + 1, \ldots, y_n) \leq$$

$$N(y_1, \ldots, x_i, \ldots, x_j, \ldots, y_n) \leq N(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n)$$

$= N(x)$. Thus $N(x) = 1$.   Similarly, $N(x) = 0$ implies $N(y) = 0$.

10. __Proposition__. Suppose positions $i_1, \ldots, i_m$ dominate respectively $j_1, \ldots, j_m$. (Several i_r and j_r may be repeated.) If $x_{i_r} - 1 \geq y_{i_r}$ and $x_{j_r} + 1 \geq y_{j_r}$ for $r = 1, 2, \ldots, m$ and $x_i \geq y_i$ otherwise, then $x \gg y$.

Proof: Apply proposition 9 m times.

Example: Suppose position 1 dominates position 2. Then $(31222222) \gg (22222222)$ and $(22222222) \gg (13222222)$. However, it can not be concluded that $(22222222) \gg (23222222)$ or $(21222222) \gg (13222222)$.

Now, if there are dominance relations, P will work as follows:
a_i has been chosen and $N(a_i)$ has been determined. T_{i-1} is known.
P is to compute T_i. First, let D_i be the set of tuples formed as
follows: If r dominates s and $N(a_i) = 1$, form

$(a_1, \ldots, a_r+1, \ldots, a_s-1, \ldots, a_n)$,

$(a_1, \ldots, a_r+2, \ldots, a_s-2, \ldots, a_n)$, etc. If r dominates s
and $N(a_i) = 0$, form $(a_1, \ldots, a_r-1, \ldots, a_s+1, \ldots, a_n)$,

$(a_1, \ldots, a_r-2, \ldots, a_s+2, \ldots, a_n)$, etc. Do this for each
dominance relation r over s. (One at a time.) D_i is the set of
all tuples constructed this way.

Examples: 4 inputs at 3 levels. (1) $a_i = (2222)$, $N(a_i) = 1$,
and 1 dominates 3. $D_i = \{(3212)\}$. (2) $a_i = (2222)$, $N(a_i) = 1$,
1 dominates 3, and 2 dominates 3. $D_i = \{ (3212), (2312) \}$.

Now $N(a_i) = N(d_j)$ for all d_j in D_i. Use P with a_i and $N(a_i)$
to get table $T_{i0} = T_i$ without dominance as in Part A above. Use
P with d_1 and $N(d_1)$ to get table T_{i1} = table computed from T_{i0} as
in Part A. Continue until finally T_{it} = table computed from
T_{it-1} as in Part A with d_t and $N(d_t)$ where d_t is the last element
of D_i. Then let $T_i = T_{it}$.

PART C

Program Q chooses a_i from table T_{i-1} as follows: (1) Let
$\underline{S} = \{ x : \min(\#_{i-1}(x), \#'_{i-1}(x)) \geq \min(\#_{i-1}(y), \#'_{i-1}(y))$ for all y $\}$
(2) Let $\overline{S} = \{ x : x$ in $\underline{S}$ and $\max(\#_{i-1}(x), \#'_{i-1}(x)) \geq$

$\max(\#_{i-1}(y), \#'_{i-1}(y))$ for all y in $\underline{S}$ $\}$

(3) a_i = one of the elements of $\overline{S}$ (arbitrary choice)

Thus a_i is chosen so that, whether it wins or loses (i.e.
whether $N(a_i) = 1$ or 0), it is sure to determine at least as many
outcomes as any other tuple, and, of all those tuples with this
property, no other can determine more.

PART D

With P and Q as described above, the graph of tuple x
vs. the number of outcomes determined by x has the form shown in figure 5.
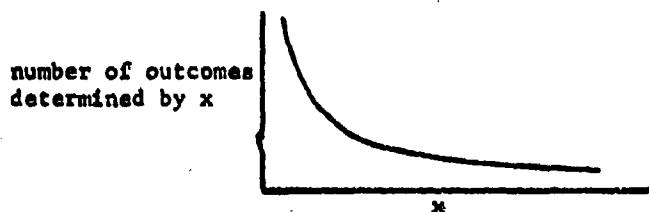
figure 5

For investigations in which there will be relatively few
actual experiments on the system (e.g. 20 out of 6561), this may
be quite satisfactory, since the low part of the curve will never
be reached. However, when more experiments are to be run (say 100
out of 6561), it will be preferable for the curve to have the
form shown in figure 6.
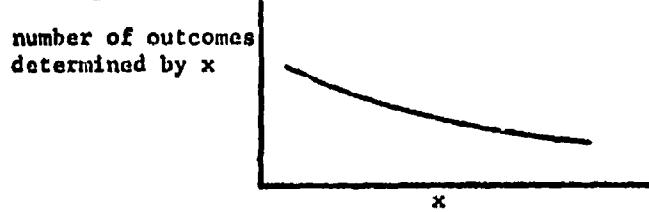
figure 6

The program $Q(w,z)$, which is a modification of Q, can be used in this situation. The parameters w and z depend on the total number of experiments possible on the system and the number of experiments which are to be run.

Parameter w is the number of outcomes that the running of one tuple x should determine. Good results have been obtained with $w = 2$ and $w = 4$. Parameter z is a number greater than 0 and less than or equal to 1. When $z = 1$, $Q(w,z) = Q$. Roughly, z indicates the slope desired in figure 2, with the slope nearer to zero as z gets small. Good results have been obtained with $z = .3$ and $z = .5$.

$Q(w,z)$ chooses a_i from table T_{i-1} as follows: (1) Using Q, choose a'_i as usual. (2) Using Q again and rejecting those tuples x such that $\min(\#_{i-1}(x), \#'_{i-1}(x)) > z \cdot \min(\#_{i-1}(a'_i), \#'_{i-1}(a'_i))$, choose a_i. (3) If $\min(\#_{i-1}(a_i), \#'_{i-1}(a_i)) < w$, then redefine a_i to be a'_i, the original choice of Q. If $\min(\#_{i-1}(a_i), \#'_{i-1}(a_i))$ is still less than w, terminate SMEDAL.

PART E

A different refinement of program Q is possible via a heuristic search algorithm. This is discussed in section V.

-345-

P AND Q FOR k > 1

We will outline the changes to be made in P and Q if k = 2. (outputs 0,1, and 2) The generalization to other cases will be analogous.

PROGRAM P

The table T_i will consist of entries of the form:

x, $\#^0_i(x)$, $\#^1_i(x)$, $\#^2_i(x)$ where $\#^v_i(x)$ = the number of outcomes determined if $N(x)$ = v for v = 0,1,2.

Given T_{i-1}, a_i, and $N(a_i)$, table T_i is computed as follows:
(1) If N(a) = 0, then $\#^0_i(x)$ = $\#^0_{i-1}(x)$ - $\#^0_{i-1}(glb(a_i,x))$, and
$\#^v_i(x)$ = $\#^v_{i-1}(x)$ for   v = 1,2.   (2) If $N(a_i)$ = 2, then
$\#^v_i(x)$ = $\#^v_{i-1}(x)$ for v = 0,1, and $\#^2_i(x)$ = $\#^2_{i-1}(x)$ - $\#^2_{i-1}(lub(a_i,x))$.
(3) Note: In this section, z >> w will mean $z_j \geq w_j$ for all j.
If N(a_i) = 1 : First, if x >> a_i, set $\#^0_i(x)$ = 0 and if x << a_i,
set $\#^2_i(x)$ = 0. Otherwise, set $\#^v_i(x)$ = $\#^v_{i-1}(x)$ for v = 0 or 2.
Second, when the above is done for all x, $\#^1_i(x)$ = the number of
tuples y such that N(y) is still undetermined and either y >> x and
$\#^2_i(y)$ = 0   or y << x and $\#^0_i(y)$ = 0.

Using the procedure described above, a tuple is determined
when two of its table entries are zero. (For the k = 1 case (section III),
a tuple was determined when one of its entries was zero.)

Concerning 3), the $N(a_i) = 1$ case, $\#^1_i(x) =$ the number of
tuples still undetermined whose outcomes would be determined if
$N(x) = 1$. Now $N(x) = 1$ gives information only about those tuples y
such that $y \gg x$ or $y \ll x$; thus, $N(x) = 1$ and $y \gg x$ implies
$N(y) \geq 1$, and $N(x) = 1$ and $y \ll x$ implies $N(y) \leq 1$. Such y are
determined only under the circumstances listed under 3) above.

PROGRAM Q

The generalization to the case $k = 2$ is immediate.

SECTION V

EXTENDING Q USING HEURISTIC SEARCH TECHNIQUES

For simplicity assume $k = 1$ throughout. Program Q chooses
a_i as the tuple that will gain at least as much as any other tuple
and will gain no less than any tuple with this property. However,
it is possible that Q will choose a_i and a_{i+1} such that each is as
above, while there exist two tuples a'_i and a'_{i+1} such that if
these were chosen, the total number of outcomes determined (over
the two steps i and $i+1$) will be more than that determined if a_i
and a_{i+1} are chosen.

For example, if in table T_{i-1} we had $\#_{i-1}(a_i) = 7$,
$\#'_{i-1}(a_i) = 7$, $\#_{i-1}(a'_i) = 6$, and $\#'_{i-1}(x) = 6$, then Q would
choose a_i over a'_i. Then it could happen that a_{i+1} would be the
best choice in T_i where $\#_i(a_{i+1}) = 3$ and $\#'_i(a_{i+1}) = 3$, while

the best choice, a'_{i+1}, in T'_i, the table formed if a'_i is chosen, would have $\#_i(a'_{i+1}) = 5$ and $\#'_i(a'_{i+1}) = 5$. Clearly, a'_i and a'_{i+1} would be a better pair of tuples to run (as they determine 11 as compared to 10 outcomes), but Q would not know this because it looks only, so to speak, one level ahead each time it makes a choice. To correct this, Q could be extended using the following heuristic search algorithm, which is patterned after that described in Slagle and Lee [4].

Assume we have table T_{i-1} , and we wish to choose a_i, the "best" tuple to run next.

A. First, we describe a single level algorithm. (Program Q) Let $MN(x)$ = minimum number of outcomes determined if x is run = $\min(\#_{i-1}(x), \#'_{i-1}(x))$, and let $MX(x)$ = maximum number of outcomes determined if x is run = $\max(\#_{i-1}(x), \#'_{i-1}(x))$. Then a_i is chosen so that $MN(a_i)$ is maximal over all x and $MX(a_i)$ is maximal over all x such that $MN(x) = MN(a_i)$.

B. Now we describe how to extend the algorithm to several levels. Refer to figure 7 on page 20.

1. Determine parameters N and M, defined below. The values for N and M will depend on such factors as the resources available and the purposes of the research.

2. Level 1: Find the best N tuples $a_{i1}, \ldots, a_{iN}$ by the single level method.

3. Level 2: For all j there are two cases: a_{ij} wins and a_{ij} loses. For each of these cases, choose best tuple b_{ijk} by the

single level method where $j = 1,2, \ldots ,N$ and $k = 0$ corresponds to a_{ij} wins, $k = 1$ corresponds to a_{ij} loses. Determine minimum number of outcomes $m_{2j1} = \#_{i-1}(a_{ij}) + MN(b_{ij1})$ and $m_{2j0} = \#'_{i-1}(a_{ij}) + MN(b_{ij0})$. Let $MN_{2j} = \min(m_{2j1}, m_{2j0})$ and $MX_{2j} = \max(m_{2j1}, m_{2j0})$. Now choose tuple a_{ij_0} where MN_{2j_0} is maximum for all j and MX_{2j_0} is maximum for all j such that $MN_{2j} = MN_{2j_0}$.

Compare MN_{2j_0} with MN_{21} where a_{11} is best single level choice. If $MN_{2j_0} - MN_{21} \geq M$, proceed to level 3. If $MN_{2j_0} - MN_{21} < M$, a_{ij_0} is chosen. (In other words, we were examining T_{i-1} to find a_i. Let $a_i = a_{ij_0}$.)

4. Level 3: Consider two cases: a_{ij_0} wins and a_{ij_0} loses. Using the single level algorithm, find the N best tuples for each case. Proceed in each case as in level 2. Use the case with the least potential yield for the test with parameter M.

5. Continue until process stops.

Notes: (1) If process does not stop until all outcomes are determined, it would be a branch and bound technique. (2) The significance of parameter N is that N branches of the decision tree will be investigated at each level. (3) The significance of parameter M is that if the level j search does not yield M more outcomes than the level j-1 search, then no further levels are investigated.

-349-
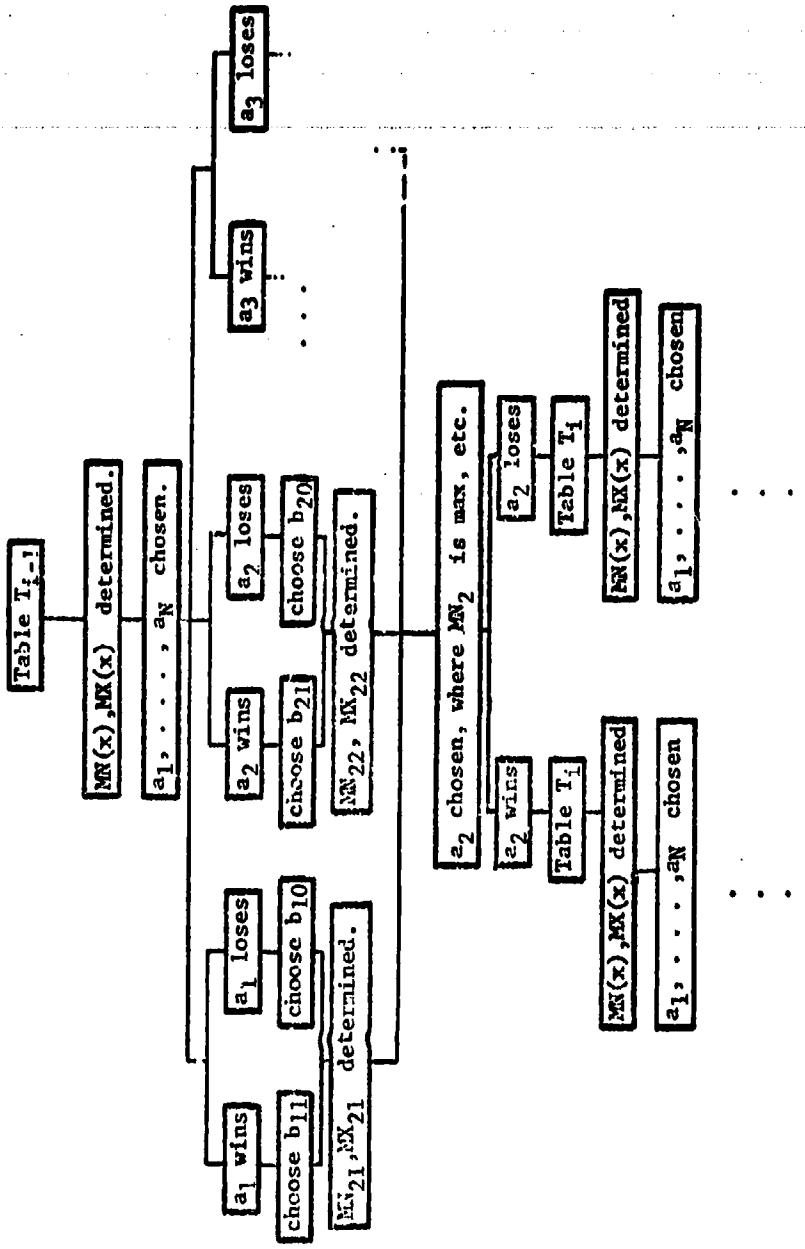
Figure 7

HEURISTIC SEARCH ALGORITHM DECISION TREE

For simplicity the subscript i has been uniformly omitted.

Also a_{1jo} in text is a_2 in diagram.

SECTION VI

OTHER APPROACHES

1. Let S be the full set of n-tuples that define our problem. S is a partially ordered set (POS) with the order relation $>>$ as defined in section III, p. 12. A subset C of S is called a _chain_ if x,y in C implies either $x << y$ or $x >> y$. A _decomposition of S_ is a family $(C_1, \ldots, C_t)$ of disjoint chains whose union is all of S. Such a decomposition is called _minimal_ if, for any other decomposition $\{K_1, \ldots, K_r\}$, $t \leq r$.

There exist various methods for finding a minimal decomposition for a POS, and these could be applied before any experiments are made on the system. (See Ivanescu [2].) If there are dominance relations, it is possible that the POS would decompose into several long chains. (Example: If position i dominates i+1 for all i, then there would be only one long chain!) Restricting our attention to such chains would allow us to "linearize" the problem and could greatly simplify the investigation.

2. A geometric approach can be formulated by replacing program Q with program G, described below. Consider the set of all undetermined tuples. (We are looking at table T_{i-1}.) These are points on a lattice in n-space. G chooses for a_i one of the tuples closest to c, the center of mass of these points. Intuitively c has the property that any n-1 dimensional subspace (a plane, if n = 3) through c has about the same number of points on one side

as the other. It is not hard to see that for any tuple x there exists an n-1 dimensional subspace through x such that the tuples determined if $N(x) = 1$ are on one side and those determined if $N(x) = 0$ are on the other. Since we wish to choose a_i in a way that tends to minimize $|\#_{i-1}(a_i) - \#'_{i-1}(a_i)|$, the points closest to c might be good candidates for optimal choices.

A few cases testing G have been computed, and these indicate that G is often close to Q but sometimes falls short. Extensive tests have not been made.

PROBABILISTIC EXTENSION OF THE GENERAL PROBLEM

It has been suggested that it would be worthwhile to extend the general problem in the following way: associate with each tuple x _a priori_ probabilities $p_0(x)$, $p_1(x)$, . . . ,$p_k(x)$ where $p_i(x)$ is the probability that $N(x) = i$. Then it would be necessary to define a program Q_p to replace Q so that optimal choices are made using the probability data. We are currently working on this aspect of the problem.

REFERENCES

1. Hunter, J.S. and T.H. Naylor. Experimental Designs for Computer Simulation Experiments. *Management Science* 16,7 (March 1970), 422-434.

2. Ivanescu, Petru L. *Pseudo-Boolean Programming and Applications*. Springer-Verlag Lecture Notes No. 9, 1965, 34-35.

3. McArthur, D.S. Increasing the Efficiency of Empirical Research. Esso Research and Engineering, PRD, Rpt. No. RL34M61, May 22, 1961.

4. Slagle, J.R. and Richard Lee. Application of Game Tree Searching to Sequential Pattern Recognition. *Communications of the ACM* 14, 2 (Feb. 1971), 103-110.

APPENDIX

SAMPLE SMEDAL OUTPUT FOR A SIMPLE CASE.

Note: N-sequence, referred to on the next page, is the
sequence of outcomes determined by experiment.
Also, the efficiency of the algorithm is considerably
improved for larger problems. (See the Table in Section 2)

SMEDAL output for case:

3 inputs at 3 levels each
2 outputs (0 and 1)
N-sequence; alternating values

A-1

RUN 1

222 HAS BEEN RUN WITH OUTCOME 0
NEXT RUN SHOULD BE 223

SUMMARY TABLE AFTER RUN 1

NUMBER FOUND FEASIBLE 0
NUMBER FOUND UNFEASIBLE 8
NUMBER UNDETERMINED 19

table T_1

111	27	0
112	18	0
113	9	1
121	18	0
122	12	0
123	6	2
131	9	1
132		2
133	3	5
211	18	0
212	12	0
213		2
221	12	0
222	8	0
223	4	4
231	6	2
232		4
233	9	10
311	6	1
312		2
313	3	5
321	6	2
322	4	4
323	2	4
331	3	10
332	2	5
333	1	10
		19

RUN 2

223 HAS BEEN RUN WITH OUTCOME 1
NEXT RUN SHOULD BE 331

SUMMARY TABLE AFTER RUN 2

NUMBER FOUND FEASIBLE 4
NUMBER FOUND UNFEASIBLE 8
NUMBER UNDETERMINED 15

table T_2

111	23	0
112	24	0
113	5	1
121	14	0
122	8	2
123		7
131		1

-354-

132	4	2
133	1	5
211	14	0
212	8	0
213	2	2
221	8	0
222	4	0
223	0	4
231	4	2
232	2	4
233	0	10
311	7	1
312	4	2
313	1	5
321	4	1
322	2	4
323	0	2
331	2	10
332	1	5
333	0	1C
		0	19

RUN 2

331 HAS BEEN RUN WITH OUTCOME 1

NEXT RUN SHOULD BE 132

SUMMARY TABLE AFTER RUN 3

NUMBER FOUND FEASIBLE	6
NUMBER FOUND UNFEASIBLE	8
NUMBER UNDETERMINED	13

Table T₃

111	21	0
112	13	0
113	5	1
121	12	0
122	7	0
123	5	2
131	3	1
132	1	5
133	12	3
211	7	0
212	2	0
213	6	2
221	3	0
222	0	4
223	2	2
231	1	4
232	0	1C
233	5	1
311	3	2
312	1	2
313	2	1
321	1	0
322	0	10
323	0	5
331	0	10
332	0	10
333	0	19

A-2

A-3

RUN 4

132 HAS BEEN RUN WITH OUTCOME 0
NEXT RUN SHOULD BE 312

SUMMARY TABLE AFTER RUN 4

NUMBER FOUND FEASIBLE 6
NUMBER FOUND UNFEASIBLE 10
NUMBER UNDETERMINED 11

111	21	0
112	13	0
113	5	1
121	12	0
122	7	0
123	2	2
131	5	0
132	3	0
133	1	3
211	12	0
212	7	0
213	2	2
221	6	0
222	3	4
223	0	
231	2	1
232	1	0
233	5	5
311	3	3
312	1	2
313	2	4
321	1	1
322	0	10
331	0	4
332	0	8
333	0	17

} table T4

RUN 5

312 HAS BEEN RUN WITH OUTCOME 0
NEXT RUN SHOULD BE 123

SUMMARY TABLE AFTER RUN 5

NUMBER FOUND FEASIBLE 6
NUMBER FOUND UNFEASIBLE 12
NUMBER UNDETERMINED 9

111	21	0
112	13	0
113	5	1
121	12	0
122	7	0
123	2	2
131	5	2
132	3	0

} table T5

FUN 6

233	1	3
211	12	0
212	7	0
213	2	2
221	6	0
222	3	4
223	0	1
231	2	2
232	1	8
233	0	0
311	5	0
312	3	3
313	1	1
321	2	2
322	1	8
323	0	3
331	0	6
332	0	15
333	0	

123 HAS BEEN RUN WITH OUTCOME 0
NEXT RUN SHOULD BE 213

SUMMARY TABLE AFTER RUN 6

NUMBER FOUND FEASIBLE 6
NUMBER FOUND UNFEASIBLE 14
NUMBER UNDETERMINED 7

111	21	0
112	13	0
113	5	0
121	12	0
122	7	0
123	2	0
131	5	0
132	3	1
133	1	0
211	12	0
212	7	1
213	2	0
221	6	0
222	3	2
223	0	1
231	2	2
232	1	6
233	0	0
311	5	0
312	3	2
313	1	1
321	2	2
322	1	6
323	0	3
331	0	0
332	0	6
333	0	13

NO FURTHER RUNS WILL NECESSARILY DETERMINE MORE THAN ONE OUTCOME

Table T6

A-4

DODGE AWARDED THE 1971 SAMUEL S. WILKS MEMORIAL MEDAL

The Presentation of the Seventeenth Samuel S. Wilks Award
Made by Frank E. Grubbs

The Samuel S. Wilks Memorial Medal Award, initiated in 1964 by the U. S. Army and American Statistical Association jointly, 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, retired industrialist, Thomasville, Georgia.

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), Jerzy Neyman of the University of California (1968), Jack Youden (1969) retired from the National Bureau of Standards, and George W. Snedecor (1970) retired from Iowa State University.

With the approval of President Churchill Eisenhart of the American Statistical Association, the Wilks Memorial Medal Committee for 1971 consisted of the following:

Professor Robert E. Bechhofer   - Cornell University
Professor William G. Cochran   - Harvard University
Dr. Fred Frishman   - Army Research Office-Washington, D.C.
Dr. Badrig Kurkjian   - US Army Materiel Command, Washington,DC.
Dr. William R. Pabst, Jr.   - Washington, D. C.
Major General Leslie E. Simon   - Retired
Dr. John W. Tukey   - Princeton University
Dr. Frank E. Grubbs, Chairman   - US Army Aberdeen Research and Development Center
                                                            Aberdeen Proving Ground, Maryland

Preceding page blank

As many conferees are aware, our process of selecting the Wilks Medalist each year turns out to be a statistically significant event, you know, screening some 25-30 nominees.

The 1971 Wilks Memorial Medalist is an international authority and a pioneer in developing sampling inspection plans and quality rating methods.

He was born in Lowell, Massachusetts in 1893 and received an S. B. Degree in Electrical Engineering from the Massachusetts Institute of Technology in 1916. After teaching Electrical Engineering for a year at MIT he joined the Western Electric Company as a development engineer in the Engineering Department. In 1924 he transferred to the newly-formed Inspection Engineering Department which became part of Bell Telephone Laboratories the next year. In that department were W. A. Shewhart, G. D. Edwards and later H. G. Romig and P. S. Olmstead whose names are also well-known in the quality control and statistical fields.

Our 1971 Wilks Medalist is a Fellow, A Shewhart Medalist and an Honorary Member of the American Society for Quality Control, A Fellow of the American Statistical Association and a Fellow of the Institute of Mathematical Statistics. From the American Society of Testing and Materials he has received the Award of Merit and an honorary membership.

He is internationally known for his published articles on sampling inspection plans and the book <u>Sampling Inspection Tables</u> which was jointly authored with Harry Romig. In addition to the single and double sampling plans presented in the book, he has published many special purpose plans. They include continuous sampling plans for conveyorized production, chain sampling plans which may be used for characteristics requiring destructive or costly tests, skip-lot sampling plans applicable to chemical physical analyses of raw materials and a cumulative results plan which is superimposed on regular acceptance sampling plans having small sample sizes.

Well known also are his contributions to the development of the Army Ordnance Standard Sampling Inspection Tables which were used during World War II for the inspection of material and his contributions to the well-known MIL-STD-105 (especially 105D issued in 1963).

Our '71 Wilks Medalist also contributed significantly to the check inspection and demerit rating plans for products used by the Bell System and prepared a quarterly Quality Report that showed quality rates for important products. This system is still the basis for judging the quality of products entering the communications network.

Throughout his career he was also a teacher: first in departmental courses and Out-of-Hours courses at BTL, then in the Communications Development Training (Kelly College) program for new BTL engineers. After his retirement from the Laboratories he became a Professor in the Graduate School at Rutgers, The State University of New Jersey.

Even his work in preparing manuals for ASTM, the Z standards for the American Standards Association and definitions in the Standards Committee of ASQC were a form of teaching. Many of his teaching techniques were also illustrated in his Edward Marburg Lecture (1954 Meeting of ASTM) entitled "Interpretation of Engineering Data: Some Observations."

It has been obvious that for some minutes now the 1971 Samuel S. Wilks Memorial Medalist is Harold F. Dodge.

THE FOLLOWING ACCEPTANCE REMARKS WERE PREPARED BY HAROLD F. DODGE
THEY WERE READ AND DISCUSSED BY LESLIE E. SIMON

Mr. Chairman and Members of the Conference:

This is, for me, a very happy occasion as I express my extreme pleasure in accepting the 1971 Samuel S. Wilks Memorial Medal which honors our long-time friend Sam Wilks and the extensive contributions he made toward the applications of mathematical statistics to problems of national interest. Those who knew him well will always be grateful for his kindness and helpfulness in translating some of our individual problems into the language of mathematical statistics. I well recall that when he looked over an early manuscript on the continuous sampling plan in 1940, he asked nicely: "Did you know that this type of series is called a 'power series'?", for he could have put the question in the more provocative form: "Don't you know that this type of series is called a 'power series'?" It is sometimes the little things that you never forget.

World War II brought out the need and opportunities for broad intensive application of quality control and statistical methods. Progress in statistical quality control (SQC) had already gone fairly far in the Bell System since its initiation in 1924. There it was definitely a team activity with close cooperation between the Quality Assurance Department of the Bell Laboratories (initially called Inspection engineering) and engineers of the installation and Manufacturing organizations of the Western Electric Company. In 1924, Shewhart invented the control chart. This, together with the development of single sampling and double sampling inspection plans and quality rating methods, led, in the late 1920's and during the 1930's, to the quite general use of these quality control methods by the installation and manufacturing groups of Western Electric. We all began to talk of 3-sigma limits, assignable causes, classification of defects, single and double sampling, and LTPD and AOQL sampling tables. And in 1941 the so-called Dodge-Romig sampling inspection tables were published. Of special note, too, was the related work done by Leslie Simon in Picatinny Arsenal starting in 1934. There in the manufacturing department he successfully introduced quality control techniques including the use of control charts for controlling the weight and uniformity of powder in various types of ammunition.

In December 1940, at the request of the War Department, the American Standards Association, using its emergency procedure appointed an Emergency Defense Committee Z1, on Quality Control, with the speaker as chairman and with Leslie Simon and Ed (W. Edwards) Deming as two of its other five members. The first two standards, Z1.1 and Z1.2, on control chart principles and techniques were published in May 1941, and the third, Z1.3 on the Control Chart Method of Controlling Quality During Production in April 1942. These standards, updated, are in force today.

But the Ordnance Department of the Army also wanted active use of quality control in the inspection of ammunition and weapons, and as a result of discussions with Bell Labs Quality Assurance Department, a new Quality Control Section was created in the Inspection Branch (later transferred to the Production Service Branch) in Washington in the spring of 1942, headed by George D. Edwards, our Director of Quality Assurance, with G. Rupert Gause of Ordnance and myself as active participants. We developed a system of acceptance quality control with new standard inspection procedures, and with assistance at the Labs, a new set of sampling tables - all of which were made available for prompt use in an Ordnance-wide quality control training program.

The sampling tables were based on the new concept of an Acceptable Quality Level (AQL), and three kinds of inspection, Normal Tightened and Reduced, with rules for switching between them, as, for example, for switching to a more exacting or Tightened sampling plan when quality runs worse than the AQL. At the outset we were literally scared before making the first presentation. We were going to talk about defectives and AQL's that implied less than full conformance with specifications. And we had been hearing about directives that in effect said such things as: a contractor "will not ship defective material." But after the first session with demonstrations of sampling with a bowl of chips and a quincunx, things seem 1 to go well.

The first QC training courses were given in 15 locations starting in Washington in August and winding up in Salt Lake City in November, 1942. Each conference lasted three days, with lectures, problem periods and homework. During this period we three learned a great deal about the cooperation and dedication of our civilian and military personnel in wartime, and I might add, about the joys of night-time travel in an upper berth. The conferences were attended by qualified nominees from over 80 Ordnance establishments. During these conferences and with the rapidly expanding applications, situations arose that indicated the need for extending the tables in a number of ways. Larger lot sizes were needed, and discussions with the Signal Corps indicated that a modified approach was desirable for complex products like radio equipments made in lots of 10, 20 or perhaps 100 each. Telephone experience had taught us that "defects-per-unit" was both theoretically and practically better than "percent defective" as a measure of quality of complex products. Work on a complete set of defects-per-unit single and double sampling tables was started and early drafts were made available in 1943 to Ordnance

and the Signal Corps. These new tables, designed to fit in statistically
with the earlier percent defective tables to form a smooth series over
the complete range of lot sizes, were used in a second series of 16
training conferences that we conducted under the auspices of the Army
Service Forces (A.S.F.) through May, June and early July 1944. This time
about half the trainees were from Ordnance, the rest distributed among the
other eight departments or corps of the A.S.F. with a few from the Navy
and Air Force. The complete set of 12 Single and Double Sampling Tables
was widely used through the rest of the war and served as a reference
standard for sampling by the technical services of the military. New
Navy tables in May 1945, using closely the same procedures, included
sequential or multiple sampling plans as well as single and double.

As reported by Sam Wilks in a 1947 paper, a nation-wide program
of short intensive quality control courses was sponsored in 1943 by
the Office of Production Research and Development (OPRD). Thirty-four
so-called "8-day wonder courses" were given in nearly 25 cities to a total
of about 2000 individuals representing 800 industrial organizations.
The texts used included the Z.1 A. S. A. Standards on Quality Control and
the Army Ordnance and A.S.F. sampling tables.

After the war, a joint committee worked out a new military standard
(105A) with sampling tables that were a compromise between the Army
and Navy wartime tables. The current MIL-STD-105D is a 1963 revision
– the result of international cooperation by an ABC (America, Britain,
Canada) working group. It has many of the basic original features,
such as AQL, Normal, Tightened and Reduced Inspection, but incorpor-
ates a number of changes including some I suggested in a 1959 Rutgers
Technical Report – one of which was a much simplified rule of
switching between Normal and Tightened inspection. MIL-STD-105D
is still widely used here and abroad.

Now there are situations where it is neither convenient nor practical,
as in conveyorized production, collect product units in lots. This
led to the development in 1940 of the continuous sampling plan, CSP-1,
which uses alternate periods of sampling and 100% inspection. This plan
was made available to the Western Electric Company and following
successful use it was published in 1943. A. C. Cohen tells of its application
with "considerable success" at Picatinny Arsenal in late 1943, citing an
example of controlling weights of bag-loaded charges of smokeless
powder.

A number of special purpose sampling plans have also been made
available. Among these are the "chain" sampling plans in which the
acceptance or rejection of a lot is based on the combined sampling
results for the current lot and one or more preceding lots, with
cumulation starting afresh after each lot rejection. Chain sampling
makes possible a sizable reduction in sample size.

Skip-lot sampling plans have been developed to provide a different but simpler form of reduced inspection, by skipping the inspection of a given proportion of lots when quality is good. The structure of these plans is based on the theory of the continuous sampling plan, CSP-1, already mentioned.

A cumulative-results plan using a cumulative results criterion (CRC), published by A. F. Cone and myself in 1964 provides a special process correction feature applicable especially when lots and samples are small. Experience with its use in quality assurance inspections by A.E.C. inspection agencies suggest that the most promising solution for the small sample problem is enforced correction of the process through the use of collective history and a cumulative-results-criterion for action.

Two special projects covered by Army contracts with the Quality Assurance Department of Bell Labs may be mentioned. The first in 1952 related to setting up Quality Control Procedures (QCP's) for controlling the ballistic quality of 105mm howitzer ammunition from a single loading line. In this project at Joliet Arsenal, the goal was large lots of uniform quality and all were pleased with the successful production of an initial large lot of over 149,000 rounds. The second project was the development in 1955-56 of an over-all quality assurance plan for the NIKE I round, with a total of 38 Quality Assurance Procedures (QAP's) covering inspections, tests, quality surveys and quality rating. The intent of this plan was to provide Ordnance with a continuing basis for satisfying itself that the quality of complete rounds was what it should be. With relatively small quantities, continuous sampling inspection plans were used. Here again we learned some new terminology, as for example, when we asked what that little motor vehicle was, like an open vat on wheels crawling across an open green, and were told it was toting "nitro" and was known as an "angel-buggy."

In conclusion, as Sam Wilks said in commenting on the use of SQC during the war: "One would have to conclude that it actually took a war to demonstrate that the methods, when applied on a wide scale, constitute an important national resource."

REMARKS ON THE ACCEPTANCE OF THE S. S. WILKS MEDAL

FOR HAROLD F. DODGE BY LESLIE E. SIMON

The work of Harold F. Dodge speaks for itself as to its importance and usefulness to the Army and to the world in general. However, as one of his many old friends and colleagues, I would like to say a few brief words about his qualities as a man. Those are qualities that have endeared him to us and have also affected his work.

Among the times that I had opportunity to observe him carefully, was when he was the chairman, and I a member of the Emergency Defense Committee on Quality Control. We were writing the American War Standards, Z1.1 Z1.2 and Z1.3, to which Dodge has referred. The work was not only crucially important, but was also done under pressure because they were needed for immediate use. Dodge's personal characteristics were in large measure responsible for the meticulous correctness of the work, which enhanced the respect and confidence in which they were held, and thereby promoted their prompt and wide-spread use. In the meetings of the committee, no question asked him or suggestion made to him was ever so simple that it did not get a courteous answer, nor so complex that he summarily dismissed it from discussion. I often marveled at his patience, consideration and generosity.

He is one of the most thorough workmen that I have ever known. It would seem to me that a task was completed, but Harold would look at it critically from every possible angle, and often bring to light conditions that were valid to the problem and had not been covered. When he felt that an assignment had been completed, it was; and only new conditions or newly discovered knowledge was likely to necessitate its revision.

It is not only his competence, depth of knowledge, and willingness to help any worthy person who asks his guidance that keeps him busy in doing gratuitous work, but his warm, out-going friendliness that makes him spontaneously say, "yes"; a quality that endeares him to all who know him.

COMPARISON OF TREATMENTS GIVEN BI-VARIATE TIME RESPONSE DATA

Pearl A. Van Natta
U. S. Army Medical Research and Nutrition Laboratory
Denver, Colorado

This paper is concerned with the question of "How may treatments be 'best' compared given bi-variate time response data?" The experimenter's plan for obtaining the data has had preliminary testing. A question which I would like the panel to keep in mind during the exposition is, "Can this data lead to another plan which would allow more meaningful comparisons?"

The following abstract of the experimental situation will give a background for the specific questions which require illumination, if not answers, by statistical analysis.

Baby chicks were used in studies designed to determine whether or not any of the compounds, histidine, histamine, and aspirin, would correct the abnormality of sulfate metabolism in bone-forming regions associated with zinc deficiency, since they correct the gross leg abnormalities similarly associated with a deficiency of the trace element. These symptoms are similar to symptoms of rheumatoid arthritis in humans. The three compounds were supplied in the diet in the following concentrations: histidine, 1.0%, histamine, .2%, and aspirin, .75%. A basal feed supplied a standard amount of sulfur. Following 4 or 5 weeks of these treatments on either a Zn-deficient (14ppm) or a Zn-adequate (94 ppm) diet, the birds were orally dosed with 10 μC_1 of $Na_2{}^{35}SO_4$. The ^{35}S content of the epiphyseal plate (EP) and the primary spongiosa (PS, from the tibia was assayed at either 6, 12, or 24 hours following isotope dosage. There were 24 random samples of chicks with sample sizes ranging from 10 to 18. ([4 drugs] x [2 Zn levels] x [3 kill-times]). The data are radioactive counts from a liquid scintillation counter which are adjusted for background and efficiency of the machine, divided by the actual milligrams of tissue counted, and called dpm/mg tissue. The experimental tissue was obtained from the end of the tibia. The EP caps the bone and has a different texture than the surrounding tissue so that the experimenter has confidence that the tissue counted is actually EP. The PS lies directly under the EP and has a similar texture to the mineralized tissue beneath it. It was more difficult to isolate. The standardization of the counts was necessary because one of the manifestations of zinc deficiency is a slow growth rate so that some groups consisted of animals with much smaller tibias and hence smaller amounts of EP and PS than other groups. An unknown quantity in this experiment is the relationship of the countable sulfur to the total sulfur in the tissue. Another unknown

quantity is whether the radioactive sulfur differences between treatments are due to changed uptake, changed excretion or both. These could be incorporated in future designs. The knowledge obtained can be useful in understanding wound healing and fracture healing.

The experimenter's questions and the results of preliminary statistical analysis ($\alpha = .05$) are:

1. Does the zinc-deficient-only diet stimulate a different sulfur response from the zinc-adequate-only diet? (Figure 1).

The preliminary data when analyzed in a univariate way for each kill-time, using the Steel-Dwass 2-sample rank statistic, show that the locations of the EP values of the two treatments differ at 6 and 12 hours but not at 24 hours and that the location of the PS values differ at 12 and 24 hours but not at 6 hours. The direction of the difference is that the Zn-adequate animals have more countable sulfur in the EP tissue at 6 and 12 hours but not at 24 hours and more in the PS tissue at 12 and 24 hours but not at 6 hours.

Figure 1 shows that the passage of time produces changes in variation in the EP and PS that is more pronounced in the Zn-adequate groups than in the Zn-deficient groups. The requirement for confidence intervals for the amount of the translation between the two populations is that the only possible difference between the distributions tested is a translocation. Hence, these were not computed.

2. Do the Zn-deficient + (a drug) diets produce responses different from the Zn-deficient-only diet? (Figure 2).

The preliminary data when analyzed in a univariate way for each kill-time using the Steel many-one rank statistic gave the information that:

   a. Aspirin stimulates more sulfur incorporation in both EP (at 6 and 12 hours) and PS (at 12 hours);

   b. Histamine stimulates more sulfur incorporation in both EP (at 6 and 12 hours) and PS (all times);

   c. Histidine stimulates more sulfur incorporation in EP (at 12 hours) and PS (at 12 and 24 hours).

3. Do the Zn-adequate + (a drug) diets produce responses different from the Zn-adequate only diet? (Figure 3).

The univariate Steel many-one rank statistic gave the information that:

a. Aspirin stimulates more sulfur incorporation in EP (all times)
but not in PS (no time).

        b. Histamine stimulates more sulfur incorporation in EP (all times)
and PS (24 hours).

        c. Histidine stimulates more sulfur incorporation at EP (6 hours)
only.

        A three-way anova (approximate) on each variable separately gave
results that Zn-level, kill-time, and drugs all have significant effects.
Specifically, for EP tissue the Zn-adequate treatment produces more
sulfur than the Zn-deficient treatment; 12 hours shows more sulfur than
6 or 24; and the treatments, no drug and histidine, give an indistinguishable
sulfur response while aspirin and histamine give an indistinguishable
sulfur response that is higher than the response elicited by the other
pair. For PS tissue, the Zn-adequate treatment also produces more
sulfur than the Zn-deficient treatment; 24 hours shows more sulfur than
6 or 12; the treatment histamine shows greater sulfur response than the
other treatments.

4. If the Zn-adequate diet produces a different sulfur response from
the Zn-deficient diet, do any of the drugs, when added to the Zn-deficient
diet, produce a pattern and level like the Zn-adequate diet? (Figure 4.)

        The one-tailed Steel many-one rank statistic was used since it
was of interest to pick out the drugs which allowed less incorporation
of sulfur than the standard. The remaining drugs, whose sulfur effects
were the same or greater than the standard, could be used for further
experimentation. The information is that:

        a. Aspirin shows no different effect for EP tissue, but allows
less sulfur to be incorporated in PS (6, 12 and 24 hours).

        b. Histamine shows no different effect for EP tissue, but allows
less sulfur in PS (12 hours).

        c. Histidine shows no different effect for EP tissue, but allows
less sulfur in PS (6 and 12 hours).

        The overall impression from these myriad tests is that histamine
is "best." However, little confidence can be placed in it, since the
whole animal has been lost in all the manipulations. A multivariate
approach might be an entry.

        A parametric multivariate approach to any of these questions
has the restriction that the variance-covariance matrices from the
24 groups cannot be considered to be similar. The variation in both
measurements varies with kill-time and drug treatment. A test
which does not require homogeneous variance-covariance matrices was

reported by Maile and Schultz at the August 1971 meetings of the
ASA in Fort Collins. The test is a randomization test and little is
known about its power. It is set up for the two-sample problem and
makes a judgment as to whether the two groups are significantly different
or not. The variables should be scaled by ranking the data within each
variable or standardizing the data. If a significant difference is
found, then the k measurements can be examined individually to determine
where the differences lie without affecting the error rate.

Would this multivariate approach provide a way to get results
for question 4? Each drug group could be tested separately with the
Zn-adequate group on a six variable basis. But then how to decide which
is best if there are competitors?

The problem is now referred to the panel for *discussion*.

BIVARIATE SULFUR COUNT PLOTTED BY KILL-TIME & DRUG TREATMENT
ZINC ADEQUATE ONLY & ZINC DEFICIENT ONLY
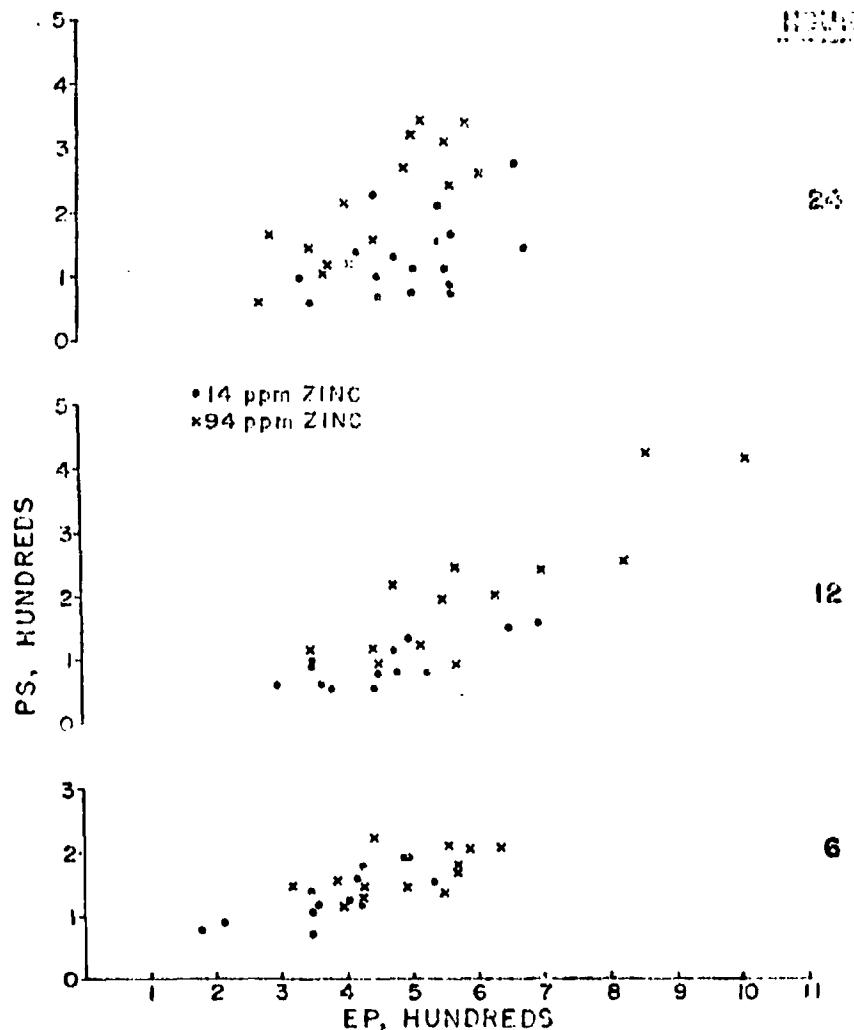
FIG. I

BIVARIATE SULFUR COUNT PLOTTED BY KILL-TIME & DRUG TREATMENT
ZINC DEFICIENT ONLY AND ZINC DEFICIENT WITH DRUGS
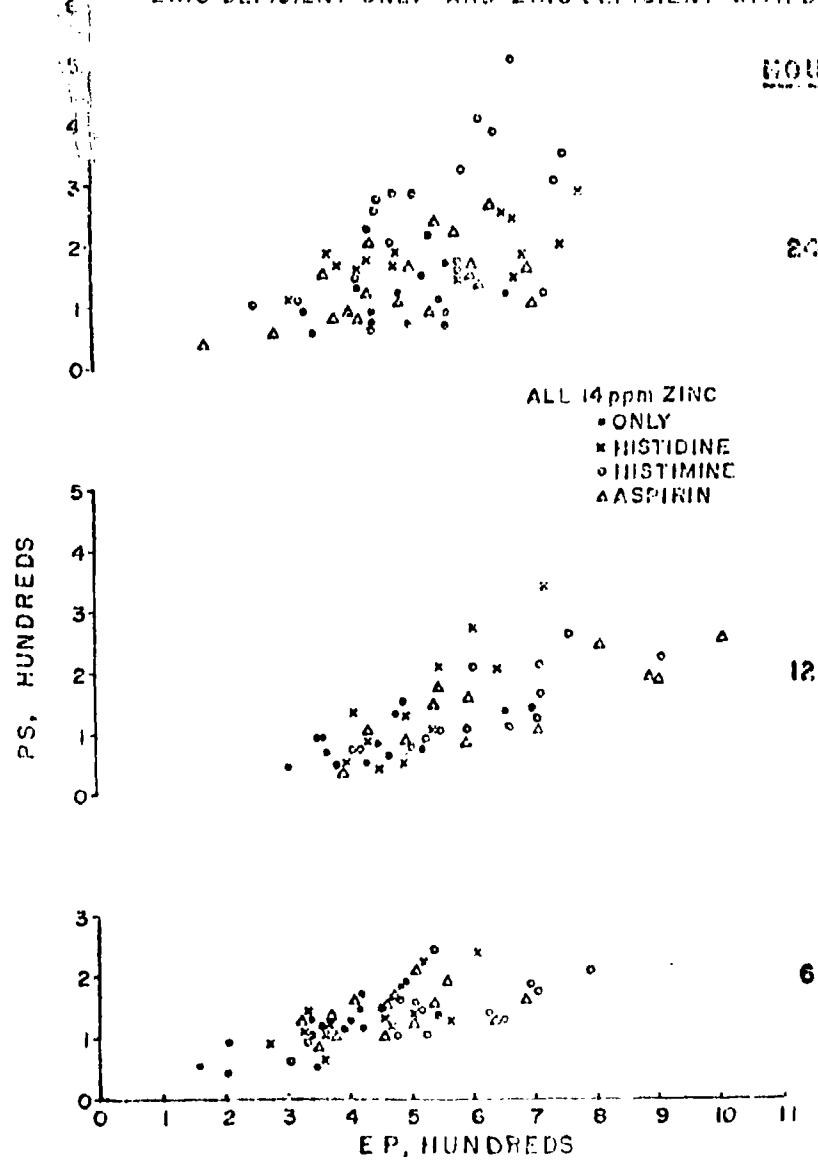
HOURS

20

ALL 14 ppm ZINC
• ONLY
× HISTIDINE
○ HISTIMINE
△ ASPIRIN

12

PS, HUNDREDS

6

E P, HUNDREDS

FIG. 2

BIVAGATE SULFUR COUNT PLOTTED BY KILL-TIME Q DRUG TREATMENT
ZINC ADEQUATE ONLY AND ZINC ADEQUATE WITH DRUGS
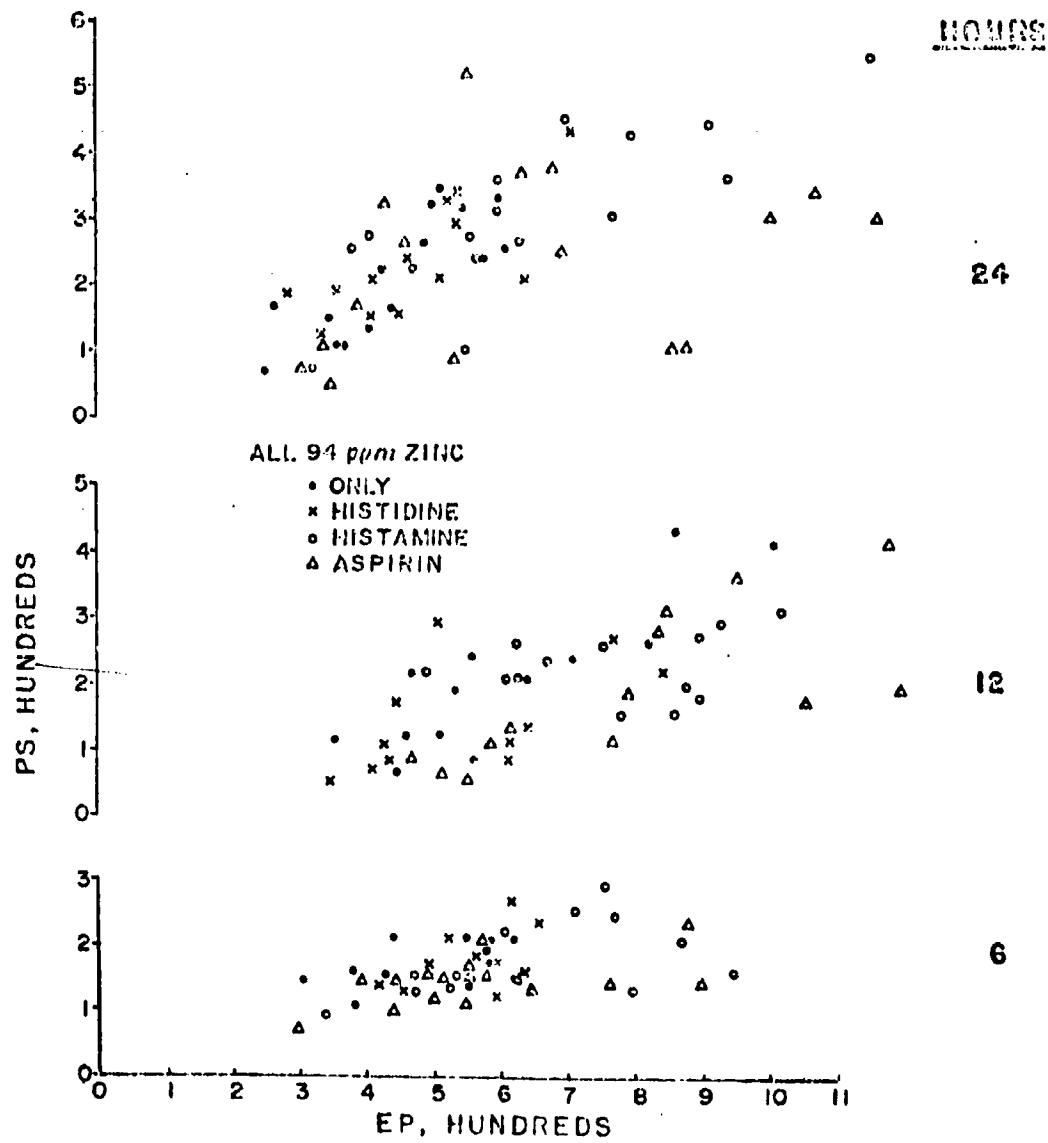
ALL 94 ppm ZINC
• ONLY
× HISTIDINE
○ HISTAMINE
△ ASPIRIN

PS, HUNDREDS

EP, HUNDREDS

FIG. 3

-373-

BIVARIATE SULFUR COUNT PLOTTED BY KILL-TIME & DRUG TREATMENT
ZINC ADEQUATE ONLY AND ZINC DEFICIENT WITH DRUGS
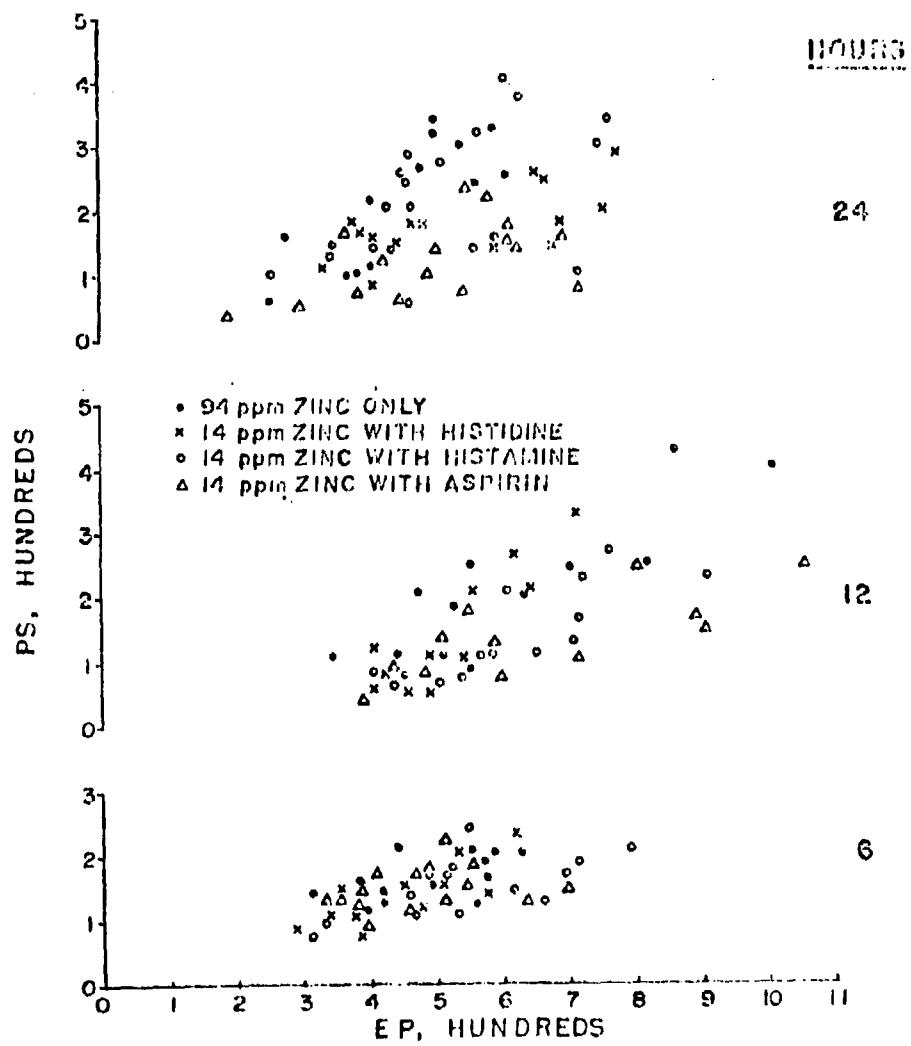
FIG. 4

EXPERIMENTAL DESIGN IN PROSPECTIVE STUDIES OF INFECTION IN MAN

Major Robin T. Vollmer
US Army Medical Research Institute of Infectious Diseases
Fort Detrick, Frederick, Maryland

The time course of an infectious disease in man is poorly understood. Especially little is known about the infection's incubation period, when no symptoms are apparent. For this reason prospective studies of infection in man have been conducted and have revealed changes in metabolism during and after the incubation period. Sandfly Fever, a self-limited febrile viral illness in man, has been the infection model, and Army volunteers have been used as test subjects.

Several biological parameters are studied throughout the course of one of these experiments. For any one parameter there are the three key variables of test subject, infection, and time which together or separately influence that parameter. Furthermore, the relationship between these parameters is also expected to be important to understanding the infectious process. With these thoughts in mind, then, the collected data has been analyzed with a linear statistical model, and some preliminary results will be presented.

The experimental design can be outlined as follows. First, let us denote a single continuous parameter being measured during the experiment with the letter "Y". It could represent the concentration of a serum substance, for example. Now we indicate two subscripts to the variable: i to represent the experimental subject and j to indicate the measurement time period. Thus the experiment with respect to one measurable parameter is represented by the array in figure 1, where time periods $1,2,\ldots,S$ comprise the baseline time interval, the period of inoculation with the infectious material being S and where subjects $1,2,\ldots,C$ are sham inoculated controls while subjects $C+1,\ldots,N$ are truly infected.

The questions to be answered by this prospective study of infection are:

1. Are there significant individual differences in the level of Y?

2. Does Y change because of infection? In particular does Y change prior to fever development?

3. Do the artificial constraints of the experiment (e.g. diet, activity level, or mental status) change the parameter Y?

An example of data from one control (marked C) and two infected (marked I) is represented in table 1. The first measurement time period is to the left of each line, and the inoculation period is the sixth. The parameter is the serum concentration of the amino acid leucine with dimensions $\mu M/100$ ml.

The next figure (figure 2) is a graph of leucine serum concentration vs. time periods of the experiment. Each point on the curve marked "infected" is an average of 9 subjects' measurements whereas the points on the "control" curve are averages of 3 subjects. Furthermore, the two points on the "infected" curve and marked with standard error brackets deviated significantly ($p < .05$) from preinfected data using a paired t test where each infected subject was compared to his own preinfected average. Such t test analysis has been the routine kind of statistical hypothesis testing performed on this data, although it was recognized there were the disadvantages of dependence between time periods and difficultly specified type I errors associated with the t test for this data. Thus a more appropriate method of analysis is sought such that it uses all the data for one parameter collected in an experiment, and we have considered first a linear statistical model of type I (see reference 1).

Referring to figure 3A the proposed linear model partitions the data matrix (Y_{ij}) into five different regions or cases. The symbol α_i represents the effect of the ith subject; β_j represents the effect of the experimental conditions on the jth day; and γ_j represents the effect of infection on day j. The error terms e_{ij} are assumed to be normally and independently distributed with equal variances and zero means. Whereas the control subjects are assumed to respond only to individual effects and artificial experimental effects, the infected subjects are assumed to respond further to the α effect of infection. The assumed model then can be represented by the matrix equation

$$(Y_{ij}) = X B + (e_{ij})$$

where (Y_{ij}), X, and B are indicated in figure 3B.

We used a version of the BMD General Linear Hypothesis program (2) and obtained the estimates in figure 4. The test of the hypothesis that $\alpha_i = 0$, $i = 1,2,...,N$ yielded an F significant at $p < .01$ so that this hypothesis could be rejected. However, the hypothesis $\beta_j = 0$ $j = 2,...,T$ and $\gamma_j = 0$, $j = S+1,...,T$ could not be rejected. It appears that this result was obtained because the linear model, unlike the t test used, considered the drop in leucine level that occurred both in infected subjects and control subjects.

This is the extent of the analysis done to date, but we currently plan to use a multivariate approach (since many parameters are measured during an experiment) to determine correlation between parameters and to determine whether the studied parameters can discriminate between infected and control states.

REFERENCES

1. Franklin A. Graybill, _Linear Statistical Models_, Vol. 1, McGraw-Hill, 1961.

2. Biomedical Computer Programs ⟨ ⟩. by W. J. Dixon, University California Press, 1971.

TABLE 1

C	16.1	14.5	19.9	21.4	20.6	15.6	16.6
	11.3	14.5	17.0	15.9	15.6	15.6	15.3
I	12.0	12.0	11.2	11.2	10.4	9.6	9.2
	4.8	6.4	9.2	9.2	10.4	8.8	12.4
I	9.2	14.2	11.0	10.4	11.6	12.6	11.8
	13.8	8.2	9.2	11.2	11.8	14.8	10.4

Figure 1

$$
\begin{array}{ccccccc}
Y_{11} & Y_{12} & \cdots & Y_{1S} & Y_{1(S+1)} & \cdots & Y_{1T} \\
Y_{21} & Y_{22} & \cdots & Y_{2S} & Y_{2(S+1)} & \cdots & Y_{2T} \\
\vdots & & & & & & \\
Y_{C1} & Y_{C2} & \cdots & Y_{CS} & Y_{C(S+1)} & \cdots & Y_{CT} \\
Y_{(C+1)1} & Y_{(C+1)2} & \cdots & Y_{(C+1)S} & Y_{(C+1)(S+1)} & \cdots & Y_{(C+1)T} \\
\vdots & & & & & & \\
Y_{N1} & Y_{N2} & \cdots & Y_{NS} & Y_{N(S+1)} & \cdots & Y_{NT}
\end{array}
$$

FIGURE 2

FIGURE 3A

$$Y_{ij} = \begin{cases} \alpha_i + e_{ij} & j = 1 \\ \alpha_i + \beta_j + e_{ij} & j = 2,\ldots,T \end{cases} \quad i = 1,\ldots,C$$

$$\phantom{Y_{ij} =} \begin{cases} \alpha_i + e_{ij} & j = 1 \\ \alpha_i + \beta_j + e_{ij} & j = 2,\ldots,S \\ \alpha_i + \beta_j + \ell_j + e_{ij} & j = S+1,\ldots,T \end{cases} \quad i = C+1,\ldots,N$$

FIGURE 3B

FIGURE 4

α	β	γ
-2.7	.5	.6
-10.6	.7	.8
2.6	1.4	-2.8
-8.4	-.4	-3.7
-5.4	-.3	-4.8
-8.3	-.2	-4.8
-3.4	-3.4	-.5
-7.4	-1.3	-1.6
-5.4	1.8	
.2	1.5	
-3.8	1.5	
	.6	
	1.5	
$F(11,135)$	$F(13,135)$	$F(8,135)$
22.5	.6	1.56

TREATMENT OF NULL RESPONSES

Genevieve L. Meyer and Ronald L. Johnson
U.S. Army Mobility Equipment Research and Development Center
Fort Belvoir, Virginia

I. INTRODUCTION.

In the design of many experiments involving equipment and human observers for data generation, a time limit or distance requirement is often present. For example, in an air observation study, the subject, serving as an air observer, has a time limitation in which to generate a response. When the subject fails to respond within the allowed time interval, his action is classified as a null response.

The problem to be considered is the anlaysis of data containing null responses where the null responses may be related to simulated artificialities and the sample size is small. Field studies are usually costly and the availability of subjects, in most cases, is limited, making re-runs prohibitive. The study of Johnson (1971) will be the subject of this paper. A similar study is that of Caviness, Maxey, and McPherson (1971).

II. DESIGN OF EXPERIMENT

The object of the Johnson field study was to evaluate the camouflage effectiveness of the two competing camouflage items. The study was conducted in the snow fields at Bridgeport, California. Two arctic-like sites were selected with snow depths of 10 to 12 feet. One item was placed at each site. The items were interchanged halfway through the test series to prevent test bias due to site peculiarities. Sixteen trained U.S. Marine air observers were used as subjects. They were flown in a Super Sky Chief Cessna 337 with all observations from the co-pilot's seat. The tester sat directly behind the subject and notified the subject of the start and termination of each run. Each subject was given an opportunity to view both test sites. Half of the subjects were flown over the one site first and half over the other site first. Two altitudes were used, 1300 feet altitude with target lateral displacement of 1000 feet, and 500 feet in altitude with target lateral displacement of 650 feet. The slant ranges were 1772 and 1032 feet, respectively. If the subjects detected the target on the higher altitude, the lower flight was cancelled because the subjects would see the camouflage items as soon as physically possible (site familiarity). The aircraft flew on a heading which minimized the effect of the sun. Each subject was debriefed at the conclusion of the test in order to obtain complete test information concerning their responses. Figure 1 represents the test design.

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

Preceding page blank

TEST MATRICES

	A_1			A_2	
	N_1	N_2		N_1	N_2
	S_5	S_5		S_5	S_5
	S_8	S_8		S_8	S_8
T_1	S_{12}	S_{12}	T_1	S_{12}	S_{12}
	S_{14}	S_{14}		S_{14}	S_{14}
	S_{16}	S_{16}		S_{16}	S_{16}
	S_2	S_2		S_2	S_2
	S_6	S_6		S_6	S_6
	S_7	S_7		S_7	S_7
T_2	S_{10}	S_{10}	T_2	S_{10}	S_{10}
	S_{13}	S_{13}		S_{13}	S_{13}
	S_{15}	S_{15}		S_{15}	S_{15}
	S_1	S_1		S_1	S_1
	S_3	S_3		S_3	S_3
T_3	S_4	S_4	T_3	S_4	S_4
	S_9	S_9		S_9	S_9
	S_{11}	S_{11}		S_{11}	S_{11}

A_1: 1300 Feet in Altitude T_1: Sunrise

A_2: 800 Feet in Altitude T_2: Noon

N_1: Camouflage Item One T_3: Sunset

N_2: Camouflage Item Two $S_1 - S_{16}$: Air Observers

Figure 1

III. RESULTS OF EXPERIMENT

The results of the experiment are as shown in Figure 2. The
response was the slant range in feet at which the detection was made.
The asterisk reflects a null response, that is, no response was made.
The few responses recorded make meaningful analysis difficult - if not
impossible.

IV. INTERPRETATION AND ANALYSIS OF RESULTS

One difficulty in the analysis of the results is the treatment of
the null responses. Presumably, detections would have occurred at
closer ranges. One approach to the problem is to assign a default
value to the null responses. Possible default values are 1700 feet for
1300 feet altitude and 1000 feet for 500 feet altitude. These default
values represent the minimum slant range for each altitude, i.e., the
"last" opportunity the subject has to make a detection before the air-
craft proceeds past ground zero. These values, obviously, overestimate
the detection probability. Yet, using zero ranges underestimates the
detection probability. Figures 3, 4, 5, and 6 show pseudo test results
using the default values and their analysis of variance, respectively.
No significant main effects or first-order interactions were found.

Another approach would be to use the available data to extrapolate
cumulative detection probability vs. lateral range and use the resulting
expected detection range as entries in the preceding test results matrix.
The results should then be better than if either of the extremal values
were used. A computer program called MATCH developed by L. Larson
at the U.S. Naval Test and Evaluation Detachment, Key West, Florida,
may be used to extrapolate these estimates. MATCH accepts a minimum
of four data points and attempts to fit these data to both a normal and
log-normal distribution. The output is an estimated curve depicting the
cumulative probability of detection along with 90% confidence intervals.
The original data in cumulative form are also plotted as a step function.
The five detections for item two were run through the MATCH program.
Since the number of detections were below the required minimum it was
not possible to use the MATCH program for item one. The results are
shown in Figures 7 and 8.

The question is what, if anything, can be said about the improvement
in estimation if these extrapolated values are used? Also, what effect
would the use of these extrapolated values have in the analysis of
variance results?

TEST RESULTS
((In Feet)

	A_1		A_2	
	N_1	N_2	N_1	N_2
	*	1772	*	5304**
	*	*	*	*
T_1	*	*	*	2453
	*	*	*	*
	*	*	*	*
	*	*	*	*
	*	*	*	*
T_2	*	*	*	*
	*	*	*	*
	*	*	*	*
	3009	*	5304**	*
	*	*	1702	*
	*	*	*	*
T_3	*	1640	*	5304**
	*	*	*	*
	*	*	*	*

*   No Response

**   Slant Range at 1 Mile

All Other Variables Same as in Figure 1

Figure 2

PSEUDO TEST RESULTS
(In Feet)

USING 1700 FEET AND 1000 FEET FOR DEFAULT VALUES

	A$_1$			A$_2$	
	N$_1$	N$_2$		N$_1$	N$_2$
	1700	1772		1000	5304**
	1700	1700		1000	1000
T$_1$	1700	1700	T$_1$	1000	2453
	1700	1700		1000	1000
	1700	1700		1000	1000
	1700	1700		1000	1000
	1700	1700		1000	1000
T$_2$	1700	1700	T$_2$	1000	1000
	1700	1700		1000	1000
	1700	1700		1000	1000
	3009	1700		5304**	1000
	1700	1700		1702	1000
	1700	1700		1000	1000
T$_3$	1700	1640	T$_3$	1000	5304**
	1700	1700		1000	1000
	1700	1700		1000	1000

Slant Range at 1 Mile = **
Minimum Slant Range at 1300 Feet Altitude = 1700 Feet
Minimum Slant Range at 500 Feet Altitude = 1000 Feet
All Other Variables as in Figure 1

Figure 3

ANALYSIS OF VARIANCE USING MINIMAL DEFAULT
VALUES OF 1700 FEET AND 1000 FEET

Source	Degrees of Freedom	Sum of Squares	F-Ratio
Altitude (A)	1	877008	1.03
Camouflage Item One and Two (N)	1	235376	.28
Time (T)	2	21187	.01
AxN	1	672465	.79
NxT	2	3666062	2.15
AxT	2	198471	.17
Error	50	47480645	--
Total	59		

$$F_{1,50,.05} = 4.03 \qquad F_{2,50,.05} = 3.18$$

Figure 4

PSEUDO TEST RESULTS
(In Feet)

USING ZEROS FOR NO RESPONSE

		A_1			A_2	
		N_1	N_2		N_1	N_2
		0	1772		0	5304**
		0	0		0	0
T_1		0	0	T_1	0	2453
		0	0		0	0
		0	0		0	0
		0	0		0	0
		0	0		0	0
T_2		0	0	T_2	0	0
		0	0		0	0
		0	0		0	0
		3009	0		5304**	0
		0	0		1702	0
		0	0		0	0
T_3		0	1640	T_3	0	5304**
		0	0		0	0
		0	0		0	0

Slant Range at 1 Mile = **

All Other Variables Same as in Figure 1

Figure 5

ANALYSIS OF VARIANCE USING ZEROS FOR NO RESPONSE

Source	Degrees of Freedom	Sum of Squares	F-Ratio
Altitude (A)	1	3103555	2.08
Camouflage Item One and Two (N)	1	695096	.46
Time (T)	2	39487	.01
AxN	1	532418	.36
NxT	2	8674222	2.90
AxT	2	390505	.13
Error	50	83066071	--
Total	59		

$$F_{1,50,.05} = 4.03 \qquad\qquad F_{2,50,.05} = 3.18$$

Figure 6

CUMULATIVE PROBABILITY OF DETECTION FOR CAMOUFLAGE ITEM TWO
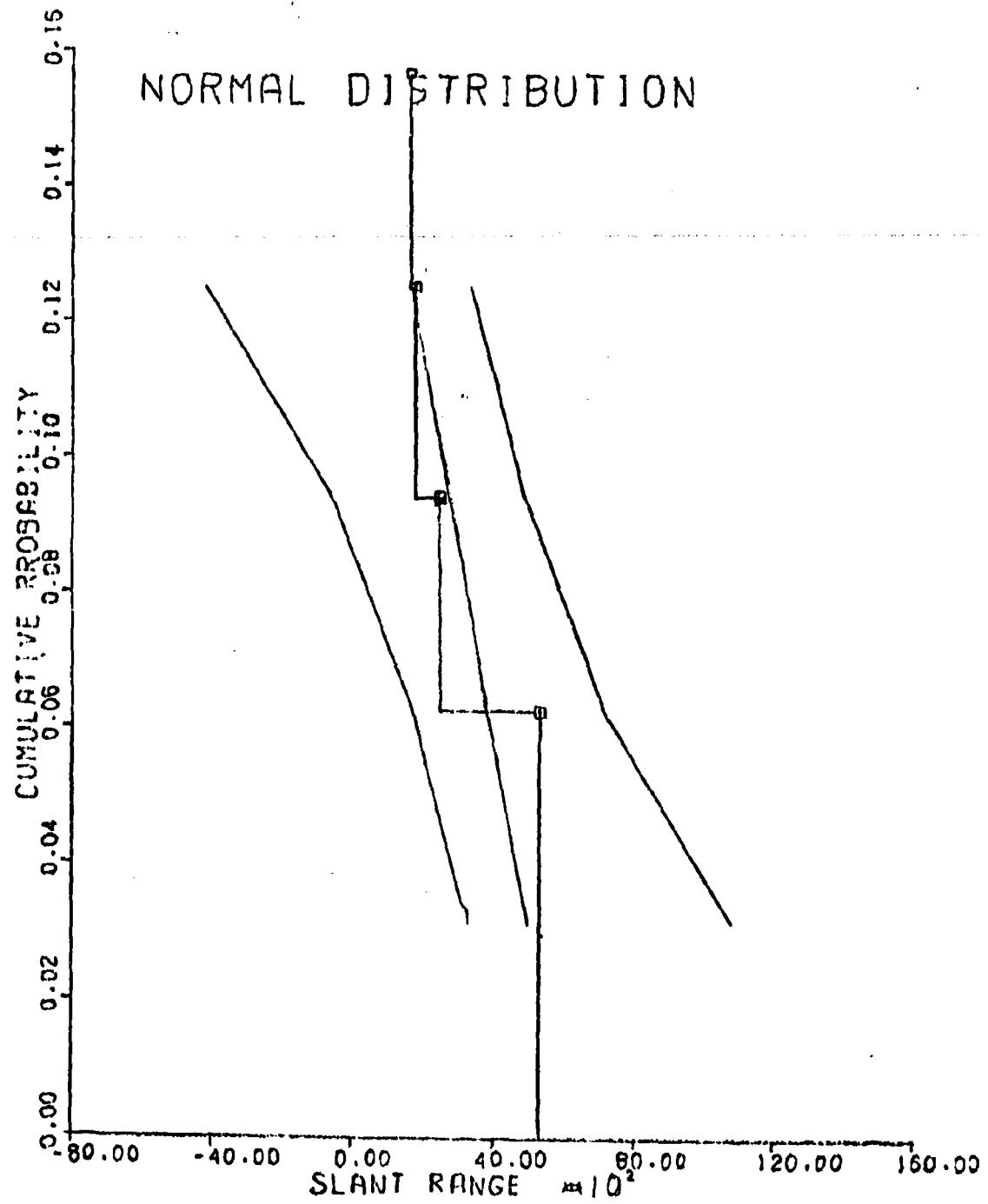
NORMAL DISTRIBUTION

Figure 7

LOG NORMAL DISTRIBUTION
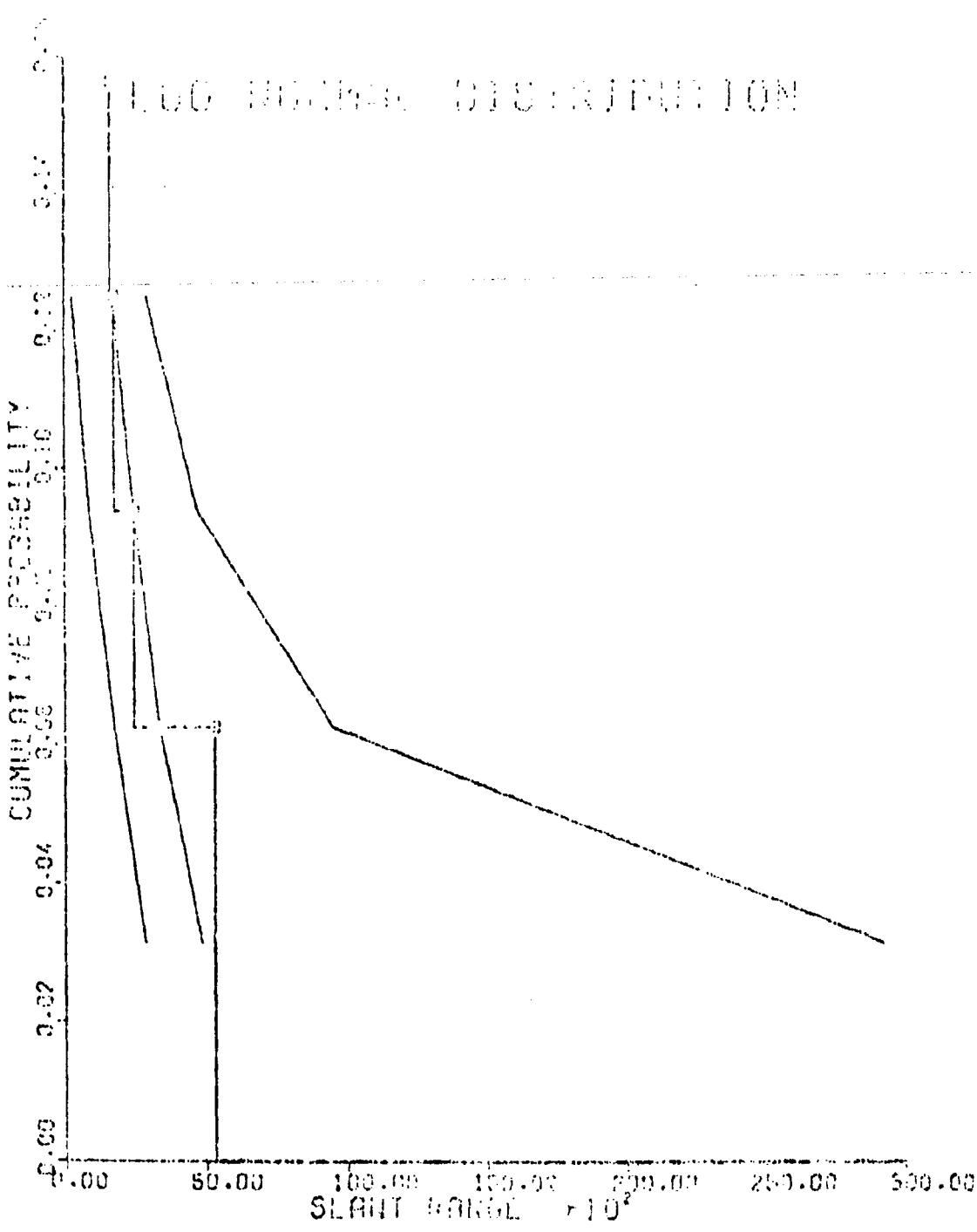
Figure 8

V. CONCLUSIONS

Difficulty is encountered in the analysis of the test data in which the null response is prevalent. A greatly increased sample size is required; however, this is not feasible for the previously stated reasons.

VI. OPEN DISCUSSION

Concepts to be considered by the panelists are as follows:

a. Improvement of the design of the experiment, considering the limited sample size.

b. Treatment of the null responses.

c. Analysis of the test data.

d. Recommendations.

REFERENCES

Cavlanis, James A., Maxey, Jeffrey L., and McPherson, James H., Target Detection and Range Estimation. Draft Technical Report, from RRO Division No. 4, Fort Benning, Georgia, 1971.

Dixon, Wilfred J. and Massey, Frank J., Jr., Introduction to Statistical Analysis. McGraw-Hill Book Co., 3rd Edition, New York, 1969.

Johnson, Ronald L., Engineer Design Test Lightweight Net Report, Arctic-Type Terrain, Woodland Terrain, and Desert Terrain. Unpublished Technical Report, U.S. Army Mobility Equipment Research and Development Center, Fort Belvoir, Virginia, 1971.

Mendenhall, W., Introduction to Linear Models and the Design and Analysis of Experiments. Wadsworth Publishing Co., Inc., Belmont, California, 1968.

COMMENTS ON THE PAPER

TREATMENT OF NULL RESPONSES
BY GENEVIEVE L. MEYER AND RONALD L. JOHNSON

James J. Filliban
National Bureau of Standards
Washington, DC

I agree with the authors Meyer and Johnson that some fundamental difficulties do exist in the design and analysis of their data, viz., e.g., 1) the duplication and hence lack of independence of some of the observations; 2) the rather small sample size; and 3) the questionable fulfillment of the ANOVA assumptions due to the multiplicity of null responses.

A general reminder about statistical testing is that for a given problem and a given sample size, there may exist a priori no possible outcome that would be statistically significant at the usual .05 and .01 significance levels. A simple example of this is to test the fairness of a coin (i.e., test H_o: P = .5) with a sample size of n = 3. Even if we get the extreme case of 3 heads (or 3 tails), the probability (1/8 ≈ 12.5%) of such occurring under H_o is significant at neither the 5% nor the 1% level.

A closely associated point to be remembered is that for a given problem and a given sample size, there may exist no outcome which is likely to occur which would be significant. For example, if the true probability of a head on a near-fair coin is P = .51, and if the sample size is 11, then the probability of a significant event at the 5% level (9 heads and 2 tails, or worse) is only .04 and so there is but 1 chance in 25 that the bias in the coin will be detected.

It is my feeling that the camouflage detection experiment of Meyer-Johnson falls into this latter category; that is to say, I suspect that the detection probabilities are so low and so close to one another that the likelihood of a significant outcome (given that a difference in the 2 camouflages does exist) is extremely low.

If I properly understand the problem then I may define a non-random variable P_i = Prob (detection using camouflage i), where i = 1,2. The overtly stated objective of the experiment is to "evaluate the camou-

flage effectiveness of 2 competing camouflage items." It is clear that the camouflage effectiveness will be (as the authors have taken into account) a function of the observer, the time of day, the altitude (or equivalently the slant range or what I shall call simply the distance), and the camouflage types (1 or 2). These 4 factors are not of equal importance. I believe that the principal problem in the analysis is that, for such a small sample size and for such apparent small *differences* in detection probabilities, the limited amount of data is being asked *too* much about too many factors and so will not likely be capable of yielding a significant result about any factor.

It is further clear that P_i is really a non-random function of the distance d and so we may speak of the function $P_i(d)$. The random element is imposed in the problem by considering a random variable $x_i(d)$ (also a function of distance) which is defined as the relative frequency of detection of camouflage i at distance d. The response model is then $x_i(d) = P_i(d) + e$ where the non-random element $P_i(d)$ is a measure of location of the random variable $x_i(d)$. We can envision the 2 probability curves as follows:

These curves have the properties of $P_i(0) = 1$, $P_i(\infty) = 0$, and $P_i'(d) \leq 0$ (monotonic decreasing). Perhaps normal, logistic, or exponential models:

$$P_i(d) = e^{(-1/2)(kd)^2}$$

$$P_i(d) = 4e^{kd}/(1+e^{kd})^2 \, , \, k>o$$

$$P_i(d) = e^{-kd}$$

can be fitted in order to approximate the non-random curves $P_i(d)$.

I believe that two separate problems are now distinguishable:
1) Which camouflage is better? (the answer to this of course depends
on our criterion of goodness of camouflage); and 2) What is $P_1(d)$ (the
probability of detection using camouflage 1) at a given distance d? The
experiment as presented is, I believe, much better suited to answer
question 2 and rather poorly suited to answer question 1.

The following specific suggestions may be of some use in regard
to the experiment as performed or to future experiments. Change the
response from many nulls and some distances (corresponding to the ques-
tion "At what distance do you detect the camouflage?") to all 0's and 1's
(for NO and YES respectively) corresponding to the question "Do you detect
the camouflage?"

The present response as given in the Test Results table reflects
too much on the question: "What does $P_1(d)$ look like at this d?" and not
enough on the null hypothesis question: "Does $P_1(d) = P_2(d)$ for this
(relevant) d?" The data in the Test Results table are effectively for
the two distances d = 1772 and d = 1032. I believe that it is proper in
this situation to split up the data (one subset for each altitude or dis-
tance) and make two separate tests of hypothesis: $H_o: P_1(d_1) = P_2(d_1)$ and $H_o:$
$P_1(d_2) = P_2(d_2)$. If the tests prove to be significant in the same direc-
tion at both distances, then this would strongly indicate that one camou-
flage is more effective than the other.

For a fixed distance, we are essentially comparing two probabilities
in which case the Sign Test (see, for example, Dixon and Massey, 2nd
edition, page 280) is an excellent way of proceeding. Complications do,
however, arise from the presence *of zeros* (ties) in the Sign Test differences.

Applying the Sign Test procedure to the high altitude (d = 1772)
data we get 2 plusses, 13 zeros, and 1 minus. It makes a difference (even
intuitively) whether there are 2 plusses, 0 zeros, and 1 minus as opposed
to 2 plusses, 10^6 zeros, and 1 minus. The relative number of zeros gives
an indication of the height of the $P_1(d)$ curves and of the relative
difference of the $P_1(d)$ curves at this d. Many zeros formed from 1-1
differences indicate that the $P_1(d)$ curves are near 1 and close to one
another. Many zeros formed from 0-0 differences indicate that the $P_1(d)$
curves are near 0 and close to one another. In the present case, the

-399-

relatively large number of zeros (13), all of which were formed from
0-0 differences, indicates that the $P_i(d)$ curves are near zero and re-
latively close for this distance d.

The presence of zeros (ties), though somewhat informative, does
invalidate the calculation of the standard binomial probabilities. To
get around this complication, I propose the following Modified Sign Test
which take into account this information on ties. As before, we are
interested in the null hypothesis H_o: $P_1(d) = P_2(d)$, where d is fixed.
After taking differences, as we would in the standard Sign Test, the
above null hypothesis is converted into: H_o: $P_+ = P_-$ (i.e. the probability
of a positive difference equals the probability of a negative difference).
However, if ties exist, then the probability of a zero difference (P_o)
also exists--let us call this ρ (i.e., $P_o = \rho$). ρ is a nuisance parameter
that is a function of the height and the relative difference of the $P_i(d)$
curves for this fixed d. Although ρ is unknown, let us momentarily treat
it as known and fixed. The probability of exactly i +'s, j o's, and k-'s
is therefore, under the *null hypothesis*, given by the trinomial expression

$$P\{(i,j,k)\} = \frac{i!j!k!}{(i+j+k)!} (P_+)^i (P_0)^j (P_-)^k$$

which may be written (in terms of ρ) as

$$P\{(i,j,k)\} = \frac{i!j!k!}{(i+j+k)!} \left(\frac{1-\rho}{2}\right)^{i+k} \rho^j$$

(since $P_o = \rho$ by definition, $P_+ = P_-$ under H_o, and $P_+ + P_o + P_- = 1$).

Using the above probability expression, we can then compute the
tail probability by simple summation over the tail region; this tail
probability will also be a function of ρ. It is now suggested that we
maximize (either analytically or numerically) this tail probability with
respect to ρ. If this maximum is still less than .05, for instance, then
a valid statistically significant result has been obtained at the 5%
level. Some minor difficulties arise in defining an appropriate tail
region due to the loss of a natural ordering in 2 or more dimensions;
however, careful consideration of various possible outcomes quickly

leads us to a realization of what constitutes a "worse event" than what was observed. Application of the above Modified Sign Test to the authors' data yields a non-significant result at the 5% level. As indicated before, this is not surprising because even though we now have the analytical tools to carry out the test, it is still felt, for this small sample size, that the magnitude and differences of the $P_1(d)$ curves are too small to make a significant result likely to occur.

A final suggestion is to change the experiment outright so as to perhaps answer more directly the question "Which camouflage is better?" This could be done by 1) physically setting up two camouflages side-by-side at each site (rather than having one camouflage at one site and the other camouflage at the other site); 2) telling the observer exactly where the sites are so that he will definitely see the two side-by-side camouflages; and 3) changing the question to be asked from "At what distance to you detect the camouflage?" to "Which of the two side-by-side camouflages do you consider more easily detectable?"

Proceeding in this manner, one can ensure the absence of null responses in which case the Sign Test is then directly applicable. We note in passing that for sample size n = 16, as is the Meyer-Johnson sample size, at least 12 of the responses would have to be of the same type for significance at the 5% level.

On the other hand, a bit more generality can be added to the testing procedure by not constraining the observer to say each and every time that camouflage 1 is definitely more conspicuous than camouflage 2 (or vice versa). It is conceivable and realistic that an observer may respond to the question "Which of the two side-by-side camouflages do you consider more easily detectable?" by "Neither looks better than the other--they both appear to be just about equally (in)conspicuous." If this type of response is permitted then the Modified Sign Test as described previously would be applicable.

REFERENCES

Bradley, J. V. (1960). <u>Distribution-Free Statistical Tests</u>. (Wright
    Air Development Division Technical Report 60-661). Washington,
    D. C.: Office of Technical Services, U. S. Department of Commerce.

Dixon, W. J. and Massey, F. J. (1957). <u>Introduction to Statistical</u>
    <u>Analysis</u> (2nd ed.). New York, N.Y.: McGraw-Hill

Noether, G. E. (1967). <u>Elements of Nonparametric Statistics</u>.
    New York, N.Y.: John Wiley and Sons, Inc.

DESIGN OF RELIABILITY EXPERIMENTS TO
YIELD MORE INFORMATION ON FAILURE CAUSES

Roland H. Rigdon
U. S. Army Weapons Command, Rock Island, Illinois

The design of reliability experiments should be changed to separate the main modes of failures. This can be done by comparing failures and variable input factors such as hours of operations, miles, rounds fired, or circuit actuations.

Each test mission or test day should be considered as a separate treatment of the factors at a high or low level. The structure is a two-level factorial. The conditional probabilities of different proportions of failures observed at high levels of the factors, given that a relationship exists between failures and the factor, can be pre-determined. This conditional probability is inverted to yield a conditional probability that a relationship exists given the failure combinations observed during the test.

The failures may be related to uptime in which case they are usually mutually independent failures and are used to predict Mean Time Between Failures (MTBF) or the failures can be used to predict Mean Miles Between Failures (MMBF) or Mean Rounds Between Failures (MRBF).

The design engineer will segregate failures into mobility and non-mobility failures. Mobility failures of the suspension and drive train are dependent upon the stresses resulting from the application of drive forces and terrain forces and may also be wear-out related.

If a non-mobility part has failures related to mileage, the design engineer knows that the failures may be due to vibration. This is especially true of a tracked vehicle because the chordal action of movement induces a longitudinal vibration of a frequency linearly related to ground speed. Thus, acceleration from minimum to maximum speed will induce an increasing frequency that will coincide at some point with the resonant frequencies of most parts.

If the failures are predicted as Mean Rounds Between Failures (MRBF) the design engineer knows that the failures may be caused by the high translation acceleration of the recoil impulse. Also, because this firing pulse is an impulse the failing part may be excited to vibrate at its resonant frequency and induce failure.

Most electrical and electronic failures are predicted as a function of uptime (MTBF). However, most failures of such equipment occur at, or are induced by, the mechanical interfaces. Since the mechanical components are in steady state during uptime and in dynamic state at energization and deenergization, many failures may be related to the number of circuit activations.

-403-

The above discussion reveals the design engineer's interest in relating failures to hours uptime, miles, rounds fired, and circuit activations. He must have this information if he is expected to redesig· a part to reduce the failure rate and achieve that redesign at a low risk.

Before discussing the proposed design of experiments, we can discuss some prior assumptions. It is assumed that the Reliability Block Diagram and the Failure Modes and Effects Analysis (FMEA) have been completed prior to designing the reliability tests. The failures are assumed to be independent. Since the lambda of a Poisson distribution would be very small, the probability of two failures on one part during one mission approaches zero. Should two or more failures occur during one mission, it is possible that the failures are not independent, i.e., a maintenance failure has occurred in the restoration of the previous failure. Or the lambda parameter of the Poisson distribution may not be small.

Most parts will have only two or three modes of failure. In real life situations, if many failures are experienced on a given part, most, if not all, will be of the same failure mode because of a weak link in the design. This consistency of failure will be obvious evidence for the design engineer to use in his redesign. The Design of Experiments herein proposed, will be most valuable in classifying a small number of failures with respect to levels of the varied test factors.

The design of present reliability tests is based upon the weapon system's mission profile with a 48 hour battlefield day as a typical mission. During this 48 hour day, the system will be operating between 18 and 24 hours and travel a specific number of miles while firing a specific number of rounds. Present practice is to establish a test mission and then perform this one mission over and over again to accumulate the desired number of hours, miles and rounds on the weapon system.

Each weapon unit should be tested over a sufficient number of missions so that the accumulated hours or miles exceeds the weapon system's expected life. If this is not possible, the accumulated interval should at least equal the expected intervals between rebuilds or overhauls. Only with this long an interval of test, can wear-out data be obtained.

The distribution function of failures can be expressed in the dimensions of hours, miles, or rounds. If all the test missions are identical, the ratios among these dimensions will remain constant. Even with some prior information that a particular failure is related to miles, the achieved failure distribution expressed in miles cannot be compared to the failure distribution expressed in hours. Therefore, there is no test of the hypothesis that a failure to miles relationship actually exists.

To overcome this dilemma and to provide data to the statistician, a new design of the reliability tests is proposed. The structure of the test design should be such that not all missions are identical. Time would remain constant over all missions, but other factors of interest would be varied to conform to a two-level factorial. As an example, a present design may test all missions running 100 miles in 24 hours. The proposed design would run half of the missions over 50 miles in 24 hours and half over 150 miles in 24 hours. The average is still 100 miles per mission. If the failures are related to time, the probability of experiencing a failure in 150 miles is equal to the probability of experiencing a failure in 50 miles, but if the failures are related to miles, the probability of failure in 150 miles is three times that in 50 miles.

The basic factorial design must be altered to prevent wear-out failures being confounded with failures related to a factor of interest. The design of experiment should be structured in such a way that at any point in a unit's life the accumulated number of missions at high and low levels should be about equal.

Appendix I contains a design for a hypothetical tank test. An initial factorial design is shown in Table I and then altered in Table II to remove the confounding of wear-out failures.

The analysis requires splitting the treatments into that half with a high level and that half with a low level of the factor. Then the number of failures occurring during high level missions can be compared to the total number of failures and indicated as F_i, where i varies from 0 to the total number of failures. If H_1 is the hypothesis that the failures are independent of the factor and H_2 is the hypothesis that the failures are dependent on the varied factor of interest, a set of conditional probabilities of F_i given H_j can be constructed as shown in Table IV of Appendix II.

However, the data from a test are the observed values of F_i while the information of interest is the conditional probability that the H_2 is true given the observed failure combination F_i. The conditional probabilities must be inverted.

The inversion of the conditional probability can be accomplished using Bayes' Theorem as follows:

$$Pr\ (H_j|F_i) = \frac{Pr\ (F_i|H_j)\ Pr\ (H_j)}{\sum_j Pr\ (F_i|H_j)\ Pr\ (H_j)}$$

Of interest in this case is the conditional probability that the hypothesis that failures are dependent on the factor, or H_2, is true given the observed combination of failures.

Prior to analysis using Bayes' Theorem, an a priori assumption of H_2 must be made. This can be established at a neutral point of .5 or can be varied dependent upon information from previous tests, or the a priori assumption can even be established from previous engineering analysis of the failures considered in the F_1 combination. This allows all available information to enter into the analysis.

Appendix II contains several tables of the conditional probability of $(H_2|F_1)$. The desired table is selected, depending upon the level of the assumed a priori probability of H_2. Using the observed combination of failures (F_1) the conditional probability that H_2 is true given the failure combination F_1 can be read directly from the tables. The statistician can then inform the design engineer of a definite probability that the failures are related or unrelated to specified factors. Appendix III contains examples of the use of different a priori assumptions.

Note that time as a factor has been held constant over all missions. If no relationship can be determined for failures during reliability analyses, the failures are assumed to be random and are customarily measured over a time base. Also, reliability must be measured on a time base to compute availability.

Another reason for using a constant time is to simplify testing. The 48 hour battlefield day of the mission profile can usually be compressed into a 32 hour test day in the test profile. This suggests the use of two 16 hour days of two 8 hour shifts each. It also affords the opportunity to conduct both day and night testing.

Examination of Appendix I, a sample test design, discloses a disadvantage of this suggested design. Under present designs only one test profile or mission must be structured for several repetitions. The new design may require as many as 32 test profiles. It is recommended that test modules of 5 miles each be designed. Different modules would have different speeds, terrains, number of rounds fired, etc. A test profile of 50 miles would then be the sum of 10 or more modules while a 150 mile profile could contain 30 or more modules, some repeated.

Whatever methods are used to structure the reliability test, the design should provide data in such a form that it can indicate both what part failed and some conditional relationships with the test environment.

APPENDIX I

HYPOTHETICAL TANK RELIABILITY TEST

Assume a mission profile for the 48-hour battlefield day as follows:

    24 hours run time
    100 miles movement (M)
    20 rounds of main armament (full charge) (R)
    120 electrical circuit actuations (A)
    200 rounds of cupola machine gun (C)
    200 rounds of co-ax machine gun (X)

The electrical circuit actuations are the energization of all intermitten'
circuits in an established sequence.

Expected life of the weapon system at overhaul is 5,000 miles so
conduct approximately 50 missions on each tank to acquire wear-out data.
Replicate with a sample of five tanks for a total of 25,000 miles.

Construct an initial factorial design as shown in Table I using the
following levels of each factor:

    Constant = 24 hours run time
           M = 50 and 150 miles
           R = 10 and 30 rounds
           A = 60 and 180 actuations
           C = 100 and 300 rounds
           X = 100 and 300 rounds

Reorder the runs of the initial factorial into the altered factorial of
Table II. This will remove the effects of wear from the analysis.

The total of 48 missions per tank will yield 24,000 miles,
5,760 hours, 4,800 main gun rounds, 28,800 actuations, 48,000 cupola gun
rounds and 48,000 coaxial gun rounds. The expected total number of
failures would range from 190 to 240 on the system. Table III shows the
different levels of a reliability block diagram with the range of the
number of blocks at each level and some ranges of expected number of
failures of a particular block to warrant investigation.

TABLE # I
INITIAL FACTORIAL DESIGN AT $(2^5)+(\frac{1}{2})(2^5)$

RUN NO.	M miles	R rounds	A actions	C rounds	X rounds	RUN NO	M MILES	R RDS	A ACTIONS	C RDS	X RDS
1	50	10	60	100	100	25	50	10	60	300	300
2	150	10	60	100	100	26	150	10	60	300	300
3	50	30	60	100	100	27	50	30	60	300	300
4	150	30	60	100	100	28	150	30	60	300	300
5	50	10	180	100	100	29	50	10	180	300	300
6	150	10	180	100	100	30	150	10	180	300	300
7	50	30	180	100	100	31	50	30	180	300	300
8	150	30	180	100	100	32	150	30	180	300	300
9	50	10	60	300	100	33	50	10	60	100	100
10	150	10	60	300	100	34	150	10	60	100	300
11	50	30	60	300	100	35	50	30	60	100	300
12	150	30	60	300	100	36	150	30	60	100	300
13	50	10	180	300	100	37	50	10	180	100	300
14	150	10	180	300	100	38	150	10	180	100	100
15	50	30	180	300	100	39	50	30	180	100	100
16	150	30	180	300	100	40	150	30	180	300	100
17	50	10	60	100	300	41	50	10	60	300	300
18	150	10	60	100	300	42	150	10	60	300	300
19	50	30	60	100	300	43	50	30	60	300	100
20	150	30	60	100	300	44	150	30	60	300	100
21	50	10	180	100	300	45	50	10	180	300	300
22	150	10	180	100	300	46	150	10	180	300	100
23	50	30	180	100	300	47	50	30	180	300	300
24	150	30	180	100	300	48	150	30	180	300	100

TABLE # II
ALTERED FACTORIAL DESIGN

Actual run	Initial run	M MILE	R RDS	A ACTS	C RDS	X RDS	Actual run	Initial run	M	R	A	C	X
1	1	50	10	60	100	100	25	37	50	10	180	100	300
2	32	150	30	180	300	300	26	44	150	30	60	300	300
3	2	150	10	60	100	100	27	38	150	10	180	100	100
4	31	50	30	180	300	300	28	43	50	30	60	300	100
5	3	50	30	60	100	100	29	39	50	30	180	100	100
6	30	150	10	180	300	300	30	42	150	10	60	300	100
7	4	150	30	60	100	100	31	40	150	30	180	100	300
8	29	50	10	180	300	300	32	41	50	10	60	300	300
9	5	50	10	60	100	100	33	9	50	10	180	100	100
10	28	150	30	180	300	300	34	24	150	30	60	300	300
11	6	150	10	60	100	100	35	10	150	10	180	100	100
12	27	50	30	180	300	300	36	23	50	30	60	300	300
13	7	50	30	60	100	100	37	11	50	30	180	100	100
14	26	150	10	180	300	300	38	22	150	10	60	300	300
15	8	150	30	60	100	100	39	12	150	30	180	100	100
16	25	50	10	60	100	100	40	21	50	10	60	300	300
17	33	50	10	60	300	300	41	13	50	10	180	100	100
18	48	150	30	180	100	100	42	20	150	30	60	300	300
19	34	150	10	60	300	300	43	14	150	10	180	100	100
20	47	50	30	180	100	100	44	19	50	30	60	300	300
21	35	50	30	60	300	300	45	15	50	30	180	100	100
22	46	150	10	180	100	100	46	18	150	10	60	300	300
23	36	150	30	60	300	300	47	16	150	30	180	100	100
24	45	50	10	180	300	100	48	17	50	10	60	100	300

TABLE III

Levels Of The Reliability Block Diagram Pyramid

Level	Name	Number of Units	Expected Threshold or Failures of Samples
1	System	1	200 (Weapon System)
2	Subsystem	5-9	80 (Turret)
3	Assembly	25-75	32 (Stabilization)
4	Sub-assembly	125-560	13 (Power Supply)
5	Component	625-4,200	5 (M.G. Set)
6	Part	3,125-30,000	2 (Diodes)

APPENDIX II

Bayesian Probabilities of Dependency

The failures of the test are analyzed to determine the probabilities that one of the following hypotheses is true (usually H_2):

H_1 = The failures are not dependent upon the factor varied

H_2 = The failures are dependent upon the factor varied

H_1 and H_2 are mutually exclusive, collectively exhaustive

Table IV lists the conditional probabilities of F_i given the hypothesis H_j is true. The symbol F_i represents the number of failures during those half of the missions treated at the factor's high level in combination with the total number of failures. Note that the combination is represented as a ratio but $F_i = 5/10 \neq 3/6 \neq 1/2$ because the total number of failures must always be considered.

Most parts will have only two or three or at the most five or six modes of failure. If there are more than ten failures, a consistency in the mode of failure can be seen which can be analyzed to determine the redesign. Therefore, the tables are only completed through ten failures.

Table IV the conditional probability of F_i, was constructed under the assumption that H_j was true. But after the test is conducted the data contains only F_i so that H_j must be determined from this data. Bayes' Theorem can be stated as:

$$Pr\ (H_j | F_i) = \frac{Pr\ (F_i | H_j)\ Pr\ (H_j)}{\sum_{j} Pr\ (F_i | H_j)\ Pr\ (H_j)}$$

$j = 1, 2$

$i = 1, 2, \ldots, n, n+1$

n = total number of failures

$Pr\ (H_1) + Pr\ (H_2) = 1.0$

The probabilities for H_j on the right side of the equation are the a priori values. With no previous knowledge, this can be assumed to be .5 as shown in Table V or they can assume other values as shown in Tables VI through XIII.

TABLE IV

Conditional Probability F_i given H_j

Pr $(F_i \mid H_j)$

$F_i = m/n = m$ failures at the factor's high level/n total failures

F_i	H_1	H_2
1/1	.5000	.7500
0/1	.5000	.2500
2/2	.2500	.5625
1/2	.5000	.3750
0/2	.2500	.0625
3/3	.1250	.4219
2/3	.3750	.4219
1/3	.3750	.1406
0/3	.1250	.0156
4/4	.0625	.3164
3/4	.2500	.4219
2/4	.3750	.2109
1/4	.2500	.0469
0/4	.0625	.0039
5/5	.0312	.2373
4/5	.1562	.3955
3/5	.3125	.2637
2/5	.3125	.0879
1/5	.1562	.0146
0/5	.0312	.0010

F_i	H_1	H_2
6/6	.0156	.1780
5/6	.0938	.3560
4/6	.2344	.2966
3/6	.3125	.1318
2/6	.2344	.0330
1/6	.0938	.0044
0/6	.0156	.0002
7/7	.0078	.1335
6/7	.0547	.3115
5/7	.1641	.3115
4/7	.2734	.1730
3/7	.2734	.0577
2/7	.1641	.0115
1/7	.0547	.0013
0/7	.0078	.0001
8/8	.0039	.1001
7/8	.0312	.2670
6/8	.1094	.3115
5/8	.2188	.2076
4/8	.2734	.0865
3/8	.2188	.0231
2/8	.1094	.0038
1/8	.0312	.0004
0/8	.0059	.0000

F_i	H_1	H_2
9/9	.0020	.0751
8/9	.0176	.2253
7/9	.0703	.3003
6/9	.1641	.2336
5/9	.2461	.1168
4/9	.2461	.0389
3/9	.1641	.0087
2/9	.0703	.0012
1/9	.0176	.0001
0/9	.0020	.0000
10/10	.0010	.0563
9/10	.0098	.1877
8/10	.0439	.2816
7/10	.1172	.2508
6/10	.2051	.1460
5/10	.2461	.0594
4/10	.2051	.0162
3/10	.1172	.0031
2/10	.0439	.0004
1/10	.0098	.0000
0/10	.0010	.0000

TABLE V

BAYESIAN Pr $(B_2|F_1)$, $B_2 = \underline{.5}$ a priori

TOTAL FAILURES

High level failures	1	2	3	4	5	6	7	8	9	10
0	.333	.200	.111	.059	.031	.015	.008	.004	.002	.001
1	.600	.429	.272	.159	.085	.046	.023	.012	.006	.003
2		.692	.529	.360	.219	.122	.066	.036	.018	.009
3			.771	.628	.457	.297	.175	.095	.049	.025
4				.835	.716	.541	.388	.240	.136	.072
5					.884	.792	.634	.487	.323	.192
6						.919	.851	.740	.587	.416
7							.945	.895	.810	.680
8								.962	.928	.865
9									.975	.951
10										.983

TABLE VI

BAYESIAN Pr $(H_2 \mid F_1)$, $H_2 = .9$ a priori

High Level Failures	TOTAL FAILURES									
	1	2	3	4	5	6	7	8	9	10
0	.818	.692	.530	.359	.220	.124	.065	.034	.017	.009
1	.931	.871	.771	.628	.456	.298	.175	.098	.051	.027
2		.953	.910	.835	.717	.558	.386	.240	.134	.074
3			.968	.938	.884	.791	.655	.487	.319	.192
4				.978	.958	.919	.851	.740	.587	.416
5					.986	.972	.945	.895	.810	.681
6						.990	.981	.962	.928	.865
7							.994	.987	.975	.950
8								.996	.991	.983
9									.997	.995
10										.998

TABLE VII

BAYESIAN Pr $(H_2|F_1)$, $H_2 = .8$ a priori

TOTAL FAILURES

High level failures	1	2	3	4	5	6	7	8	9	10
0	.667	.500	.333	.200	.111	.059	.031	.015	.008	.004
1	.859	.750	.600	.428	.272	.159	.085	.046	.023	.012
2		.900	.818	.692	.529	.360	.219	.122	.066	.036
3			.941	.871	.771	.628	.457	.297	.175	.095
4				.953	.910	.835	.716	.541	.388	.240
5					.968	.938	.884	.792	.634	.487
6						.978	.958	.919	.851	.740
7							.987	.972	.945	.895
8								.990	.981	.962
9									.994	.987
10										.996

TABLE VIII

BAYESIAN Pr $(H_2|F_1)$, $H_2 = .7$ a priori

TOTAL FAILURES

High Level Failures	1	2	3	4	5	6	7	8	9	10
0	.538	.368	.225	.129	.067	.036	.018	.009	.005	.002
1	.778	.636	.467	.305	.279	.100	.053	.027	.014	.007
2		.840	.724	.567	.396	.247	.141	.076	.039	.020
3			.887	.797	.663	.504	.330	.196	.110	.058
4				.922	.855	.747	.596	.424	.269	.154
5					.947	.898	.816	.689	.519	.357
6						.962	.930	.869	.769	.624
7							.975	.952	.909	.832
8								.984	.968	.938
9									.989	.978
10										.993

TABLE IX

BAYESIAN Pr $(H_2|F_1)$, $H_2 = \underline{.6}$ a priori

TOTAL FAILURES

High Level Failures	1	2	3	4	5	6	7	8	9	10
0	.429	.273	.157	.085	.046	.023	.012	.006	.003	.002
1	.692	.529	.361	.219	.122	.066	.036	.018	.009	.005
2		.772	.628	.457	.297	.175	.095	.049	.025	.013
3			.835	.716	.541	.388	.240	.136	.072	.038
4				.884	.792	.634	.487	.323	.192	.107
5					.919	.851	.740	.577	.416	.270
6						.945	.895	.810	.680	.518
7							.962	.928	.865	.761
8								.975	.951	.906
9									.983	.966
10										.989

TABLE X

°BAYESIAN Pr $(H_2|F_1)$, $H_2 = .4$ a priori

High Level Failures	1	2	3	4	5	6	7	8	9	10
0	.250	.143	.077	.040	.020	.010	.005	.003	.001	.000
1	.500	.333	.200	.111	.059	.031	.015	.008	.004	.002
2		.600	.428	.272	.159	.085	.046	.023	.012	.006
3			.592	.530	.360	.219	.122	.066	.036	.018
4				.771	.628	.457	.297	.175	.095	.049
5					.835	.716	.541	.388	.240	.136
6						.884	.792	.634	.487	.323
7							.919	.851	.740	.597
8								.945	.895	.810
9									.962	.928
10										.975

TOTAL FAILURES

TABLE XI

BAYESIAN Pr $(H_2|F_1)$, $H_2 = .3$ _a priori_

TOTAL FAILURES

High Level Failures	1	2	3	4	5	6	7	8	9	10
0	.176	.097	.051	.026	.013	.007	.003	.002	.001	.000
1	.391	.243	.139	.076	.039	.020	.010	.005	.002	.001
2		.491	.325	.193	.106	.057	.031	.016	.008	.004
3			.592	.420	.244	.152	.083	.043	.022	.011
4				.634	.521	.352	.232	.119	.063	.033
5					.765	.596	.449	.288	.169	.092
6						.830	.709	.549	.380	.235
7							.880	.785	.550	.476
8								.917	.846	.733
9									.943	.892
10										.961

TABLE XII

BAYESIAN Pr $(H_2|F_1)$, $H_2 = .2$ a priori

TOTAL FAILURES

High Level Failures	1	2	3	4	5	6	7	8	9	10
0	.111	.060	.030	.015	.008	.004	.002	.001	.000	.000
1	.272	.158	.086	.046	.023	.012	.006	.003	.001	.000
2		.360	.220	.122	.066	.036	.018	.009	.005	.002
3			.457	.297	.175	.095	.049	.025	.012	.006
4				.541	.388	.240	.136	.072	.038	.020
5					.634	.487	.323	.192	.106	.063
6						.640	.587	.416	.262	.151
7							.810	.680	.517	.346
8								.865	.763	.616
9									.907	.828
10										.935

TABLE XIII

BAYESIAN Pr $(H_2|F_1)$, $H_2 = \underline{.1}$ a priori

TOTAL FAILURES

High Level Failure	1	2	3	4	5	6	7	8	9	10
0	.052	.027	.014	.007	.003	.002	.001	.000	.000	.000
1	.142	.077	.040	.020	.010	.005	.002	.001	.001	.000
2		.200	.111	.059	.031	.015	.008	.004	.002	.001
3			.272	.159	.085	.046	.023	.012	.006	.003
4				.360	.219	.122	.066	.036	.018	.009
5					.457	.297	.175	.095	.049	.025
6						.541	.388	.240	.136	.072
7							.634	.487	.323	.192
8								.740	.587	.416
9									.810	.680
10										.865

APPENDIX III - Examples

Example I.

Assume a tank weapon system is being tested as described in Appendix I. During the 24,000 miles, 5,750 hours, and 4,800 gun rounds a total of five failures was experienced on road wheel arms.

Because this is the first test under this method of testing, no prior assumptions are made about relationships between failures and miles, rounds, or hours. Therefore, the a priori probabilities are considered neutral or .5.

One of the failures of the road wheel arm occurred during a firing exercise. Another failure was detected immediately after a firing exercise. No information is known about the other three failures. All existing evidence therefore indicates a relationship between failures and rounds fired.

The a priori probability of .5 can now be altered. Assume a probability weighting of 1.0 for the two failures related to rounds fired and a weighting of .5 for the other three. Thus, an a priori probability of .7 (3.5/5.0) is established.

The total failures occurring during missions of 30 rounds firing (high level of rounds) is found to be only two of the total of five. Inspection over the Table VIII matrix of the column under five total failures, two of which are at the high level, indicates a conditional probability of .396 that the failures are related to rounds fired, given that two of the five occur at the high level of rounds during a mission, and given an a priori probability of .7.

The relationship of failures to miles is now examined. A total of four of the five failures are found at the high level of miles per mission. Using an a priori probability of .5, Table V yields a conditional probability of .716 that the failures are related to miles (conditional probability of .284 that the failures are related to time).

This statistical evidence thus indicates that the road wheel arm failures should be measured as Mean Miles Between Failures (MMBF) despite the fact that the only available engineering evidence indicates measurement as Mean Rounds Between Failures (MRBF).

Failure due to rounds, or impulse loading, may require an entirely different fix than would a failure due to miles, or cyclic loading.

Example II.

Assume the tank weapon system of Example I was the first to be tested under this new procedure. The second weapon system to be tested is an armored personnel carrier. The road wheel arms are of a similar design and construct and have a similar strength to stress ratio. Therefore, the results of the tank test are used to establish an a priori probability of .7 that failures are related to miles and an a priori probability of .4 that failures are related to rounds.

A total of four road wheel arm failures are experienced with three at the high level of miles. Table VIII indicates a conditional probability of failure relationship to miles of .797. Two of the four failures occurred at the high level of rounds and Table X indicates a conditional probability of failure relationship to rounds of .272.

The third weapon system to be tested is a reconnaissance vehicle about the weight of the armored personnel carrier but with a main gun firing impulse nearer that of the tank. The road wheel arms are designed stronger to take this impulse.

The results of the second test are used to establish a priori probabilities of .8 and .3. A total of eight failures are recorded of which five are at the high level of miles and six are at the high level of rounds. These yield conditional probabilities of .792 for a relationship to miles (Table VII) and .549 for a relationship with rounds (Table XI).

This is counter to all previous evidence. Disregarding the previous evidence and re-establishing a neutral a priori probability of .5, Table V yields conditional probabilities of .487 and .740 for the failure relationship with miles and rounds respectively. This leads to accepting the failures as related to rounds with a conditional probability of .740. The conditional probability of relationship to miles is nearly neutral, so previous data can be accepted and the conditional probability of failure relationship to miles is stated as .792. Thus, any redesign to reduce failures should examine both MMBF and MRBF for causes.

Example III.

During testing of the tank in Example I, a "black box" in the turret has experienced ten failures. The box contains five electrical circuits and due to lack of circuit protection each failure of the box has resulted in the failure of from two to all five circuits. As a consequence, the initial failure mode cannot be isolated from its effects on other circuits.

An accelerometer has been used to measure the vibration at the surface of the "black box". Resonant frequency was 30 H_z at a ground speed of 12 MPH and with .5g lateral, 1.0g longitudinal and 3.0g vertical amplitudes. The transient loading from gun fire shock was 2.0g over 10 milliseconds.

From the above data, a tentative decision is made to isolate the box in the vertical direction. This decision is based upon the only data available under standard test procedures.

However, the new procedure allows an evaluation of the relationship between failures and miles (vibration) and rounds (gun shock). A neutral assumption of a priori probabilities of .5 is chosen. Five of the ten failures were recorded at the high level of miles and eight were recorded at the high level of rounds per mission.

Using these values to enter Table V, the matrix yields conditional probabilities of .192 and .865 for the relationships of failures to miles and rounds respectively. This is a strong indication that the gun shock transient is forcing the components inside the box into free resonant vibration. Examine each component to determine which ones can vibrate or displace under the unidirectional impulse of gun shock. A solution may be to rotate some parts 90 degrees to move the forcing function out of their vibrational axes.

Note that in this example, the tentative solution of isolation in the vertical direction may not reduce failures and may actually increase failures.

MAXIMUM LIKELIHOOD APPROXIMATION FOR GUMBEL'S LAW AND APPLICATION TO UPPER AIR EXTREME VALUES

Oskar Essenwanger
Physical Sciences Directorate
Directorate for Research, Development, Engineering
and Missile Systems Laboratory
U. S. Army Missile Command
Redstone Arsenal, Alabama

ABSTRACT. Two types of extreme value distributions are commonly applied and have been employed to fit extreme values of upper air data of wind speed, temperature, and density. Gumbel's (or Fisher-Tippett I) distribution fits well in general, while the Fisher-Tippett II distribution displays significant deviations of the observed data from the analytical curve for temperature and density when judged by the Kolmogorov-Smirnov test.

Gumbel's law requires two parameters to be estimated from the observed data x_i. Moments estimators for α and u are readily available, but not very efficient. The maximum likelihood estimators $\hat{\alpha}$ and $\hat{u}$ can be computed by iteration methods, but usually not without electronic data processing. Lieblein has introduced minimum variance linear order statistics estimators, which need table values (weighting factors of the observations) and necessitate keeping the data in sequence of recording (i.e. time).

The author has developed and tested two methods from which maximum likelihood estimators can be obtained by modification of the likelihood equations, but based on mere ranking of data. The $\hat{\alpha}$ is computed from an analytical expression rather than the lengthy iteration process ordinarily necessary to solve the likelihood equations. The approximation of $\hat{\alpha}$ and u_x by the two methods takes a minute fraction of the computer time compared with the time for the iteration of the maximum likelihood equations. Although the second method leads to a quadratic equation, one solution can readily be discarded by consideration of $\hat{u}$.

The methods were tested against the results of the maximum likelihood solution. It was found that for all three atmospheric parameters (wind speed, temperature, density) at altitude levels 1 km through 24 km (in 1 km steps) for distributions of summer and winter data the approximation methods provided excellent agreement with the true $\hat{\alpha}$ and $\hat{u}$ estimators.

1. INTRODUCTION. A variety of factors must be considered in the assessment of the atmospheric influence upon missiles and rockets. It is self-evident that extreme values of atmospheric parameters weigh heavily in the analysis of the atmospheric effect. Of great importance to the engineer is therefore the proper representation of the distribution of extreme values, especially of upper air data.

Two types of extreme value distribution would mainly be applicable to fit upper air data of wind speed, temperature and density, namely Fisher-Tippett I and II (1928). Although other frequency distributions could be employed, these two types appear best suited to represent the above mentioned data. The Fisher-Tippett I distribution has also been derived by Gumbel (1958) as the limiting distribution of the m-th distribution, and is often referred to as Gumbel's distribution. As later demonstrated, Gumbel's curve fits well in general, while the Fisher-Tippett II distribution displays significant deviation of the observated data from the analytical curve at various altitude levels. The significance was tested by the Kolmogorov-Smirnov test (see Kolmogorov 1933, and Smirnov, 1948).

From a theoretical point of view in statistical analysis, maximum likelihood estimators would be best for determination of the parameters from the observed data. These estimators cannot be expressed in explicit form, however. Iteration methods are usually applied to solve for the estimators. These computations can be expensive when a large number of samples is involved. Although Lieblein (1954) has introduced minimum variance linear order statistics estimators, his method requires that the observations are kept in sequence (e.g. time). A further disadvantage for computer use of the latter is the determination of the length of subsections. The author has therefore attempted to derive estimators which are based on the maximum likelihood equations and are close approximations to the maximum likelihood estimators. In essence for $N \to \infty$ they would be identical with them. Since the computation of these estimators does not require iteration, sequence of data or determination of adequate subsections, the deficiencies of the existing methods in computer applications are resolved. It is proven with the data samples on hand for upper air data that the estimators are in excellent agreement with the true maximum likelihood estimators, and the computation is less costly than with any other method except the moments estimators. The derivation and detailed results are presented in the following sections.

2. ESTIMATION OF GUMBEL'S DISTRIBUTION. The Gumbel extreme value distribution can generally be written in the form of the cumulative distribution

$$F(z) = \exp [- \exp (-z)] \tag{1}$$

with the normalized variable z related to the observed variate x by

$$z = \alpha (x - u). \tag{2}$$

The two constants α and u are scale and reference parameters, respectively, and must be determined from the observations by estimation procedures. Several methods to derive estimators are available.

a. Moments Estimators

It is easy to prove that

$$\alpha = \sigma_z/\sigma_x \tag{3}$$

with $\sigma_z = 1.28254$ and $s_x \to \sigma_x$, the usual moments estimator for the standard deviation. Hence the moments estimator is

$$\alpha_M = 1.28254/s_x. \tag{3a}$$

Further

$$u = \bar{x} \mp \bar{z}/\alpha \tag{4}$$

$$u = \bar{x} \mp \bar{z} \cdot \sigma_x/\sigma_z \tag{4a}$$

$$u = \bar{x} \mp 0.45005 \, \sigma_x. \tag{4b}$$

As usual, the mean value is replaced by the moments estimator, the empirical mean $x_m \to \bar{x}$.

$$u_M = x_m \mp 0.45005 \, s_x \tag{4c}$$

The top sign is valid for maxima, the lower sign for minima.
The moments estimators are simple to compute and have been employed by Gumbel (1958) most of the time, probably because at that time the computer methods were not as widespread as today.

b. Maximum Likelihood Estimators

The maximum likelihood estimators must be determined from the maximum likelihood equations, which can be readily derived (e.g. Gumbel 1958), as follows

$$1 - \hat{\alpha} \, \hat{\bar{x}} + \hat{\alpha} \, e^{\hat{\alpha}\hat{u}} (\Sigma x z^{-\hat{\alpha}x})/N = 0 \tag{5a}$$

$$1 - e^{\hat{\alpha}\hat{u}} (\Sigma e^{-\hat{\alpha}x})/N = 0. \tag{5b}$$

The mean value can be taken from

$$\hat{u} = \bar{x} - \bar{z}/\hat{\alpha} \tag{5c}$$

which is equation (4) modified for the maximum likelihood estimators. It is evident that no explicit solution can be obtained in the form of an analytical expression for $\hat{\alpha}$, although equations 5a and b could be combined and $\hat{u}$ eliminated. This would provide one equation, namely

$$1 - \hat{\alpha} \, \bar{x} + \hat{\alpha} \, (\Sigma e^{-\hat{\alpha}x})^{-1} \cdot (\Sigma x e^{-\alpha x}) = 0. \tag{5d}$$

It should be noted however, that either we assume that $\bar{x}$ can be replaced by the mean value x_m or we introduce the $\hat{u}$ back via equation (5c).

Then equation (5d) is again one equation with two parameters to be determined. Equation (5d) can be solved by iterative procedures irrespective of our choice for $\bar{x}$. Electronic data processing has made it possible today to obtain solutions readily. Equation (5d) has more than one solution, however, and a straight iteration with the moments estimators as initial conditions may not always lead to a unique solution, as it is unknown whether the maximum likelihood parameters are smaller or larger than the moments estimators. In the case that the extended forms with arbitrary $\hat{u}$ is employed, determination of $\hat{a}$ would be even more complex as the $\hat{u}$ and $\hat{a}$ pair must be checked independently, e.g., by equations (5c) or (5a). Then again, the $\bar{x}$ must be known, and we can replace $\bar{x}$ by the empirical mean from the beginning. An independent check would be the median or percentiles.

By and large, the computer costs for maximum likelihood approximations are reasonable if only a few samples are studied. In our case with 12 months, 25 altitude levels and several stations selected from characteristic climatic regions cost reduction is a substantial contribution.

c. Minimum Variance Estimators

Other deficiencies such as the determination of expected values of the estimators u and α for statistical judgment, their probability distribution, the efficiency and other properties have lead Lieblein (1954) to investigate whether optimum estimators can be found. His approach is the development of an order-statistics, largely for small samples, but later extended to larger samples. This order statistics provides an unbiased estimator, whose efficiency can be simply and accurately evaluated. We can write

$$u^* = \sum_{j=1}^{m} a_j s_j / k \tag{6a}$$

$$\beta^* = \sum_{j=1}^{m} b_j s_j / k \tag{6b}$$

where $\alpha^* = 1/\beta^*$ and a_j and b_j are weighting factors given in table form up to m = 6, by Lieblein. The s_j is the column sum of a data sample matrix, broken into $m \cdot k = N$ elements.

Present availability of electronic data processing simplifies the task of calculating u^* and β^*, and computations are less costly than the maximum likelihood estimators. Some difficulty arises in the arrangement of an $m \cdot k$ matrix, as prime numbers cannot be broken into $m \cdot k$ factors where $m \leq 6$ and m and k are integers. Furthermore, the m should always be selected as large as possible.

When $N = (m \cdot k + m')$, where m, k, $m' \neq 1$, but m and $m' \leq 6$, then

$$u^* = (w_1 u_1^* + w_2 u_2^*)/N \qquad (6c)$$

$$\beta^* = (w_1 \beta_1^* + w_2 \beta_2^*)/N \qquad (6d)$$

The $w_1 = m \cdot k$ and $w_2 = m'$, with subscript 1 denoting the computation of an estimator from the $m \cdot k$ data and subscript 2 the one from m'. Although the procedure to break N into $m \cdot k + m'$ sections can be computerized for electronic data processing, it may sometimes be an inefficient part of the program. The optimum choice of $m \cdot k + m'$ may not always be accomplished. However, the requirements to establish an $m \cdot k$ data matrix lead to other restrictions, too. First, we need the data kept in sampling order for the establishment of the m sections. This eliminates application to any sample where grouping has been made or where the sequence of sampling is unknown. The second problem is the re-arrangement into the data matrix form, as the N may not be known a priori, especially when the extreme value sample must be extracted from a larger collection of data, whose period of records differs in length and cannot be normalized a priori. These technical problems have nothing to do with the statistical background under which Lieblein has developed the solution of the minimum variance unbiased linear estimators. Under the goals set by him these estimators are the optimum solution. Again, for a few data samples with small number of observations these estimators may be the ideal answers, especially when one is interested in the computation of the efficiency of the estimators. Although computation time is reduced, the gain is not very substantial.

d) Modified Maximum Likelihood Estimators

The question arises whether approximations to the likelihood equations exist which lead to simplification, and give an explicit answer for α at small computer costs. A positive answer to the question can be found if we relate the maximum likelihood equations to the z systems since the F(z) is known. We replace therefore

$$e^{\alpha u} \, e^{-\alpha x} = e^{-z}, \qquad (7)$$

$$\alpha \, \bar{x} = \alpha \, u + \bar{z}, \qquad (7a)$$

and $\qquad \alpha \, x = z + \alpha \, \bar{x} - \bar{z}. \qquad (7b)$

Furthermore we may set $N = \Sigma\ F(x) = 1$ and obtain

$$\alpha\ u = \ell n\ \Sigma\ e^{-\alpha x} \tag{7c}$$

This leads from the original eqn. 5a to the following equation

$$1 - \hat{\alpha}_1\ \bar{x} - \hat{\alpha}_1\ \Sigma\ x\ e^{-z} = 0 \tag{8}$$

or $\qquad \hat{\alpha}_1\ [\bar{x} - \Sigma\ x\ e^{-z}] = 1.$ $\qquad\qquad$ (8a)

We may replace $\bar{x}$ by the empirical mean x_m. The term in the summation can be computed by substituting for the population $F(z)$ the empirical cummulative $F(x)$. Then

$$F(x) = \exp\ [\ - \exp(-z)] \tag{8b}$$

and $\qquad z = \ell n\ [- \ell n\ F(x)].$ $\qquad\qquad$ (8c)

Since for $N \to \infty$ the $F(x) \to F(z)$, the maximum likelihood estimator $\hat{\alpha}$ and the $\hat{\alpha}_1$ (approximation from equation 8a) become identical.

From the original maximum likelihood equation we can further develop a second solution by multiplication of the maximum likelihood equation with x. After some arithmetic operation we derive

$$\hat{\alpha}_2^2\ (\Sigma\ x^2\ e^{-z} - \Sigma\ x^2) + 2\ \hat{\alpha}\ \bar{x} - \bar{x} = 0. \tag{9}$$

This is a quadratic equation and has two solutions, which lead to a positive and negative u. The answer with the negative u can be discarded.

3. COMPARISON OF THE ESTIMATORS

The first task would be estimating the error which is introduced by replacing $F(z)$ by $F(x)$. Assume that we know x to a degree $x + \varepsilon_x$, which gives an error in z of $z + \varepsilon_z$. Then

$$N\ \varepsilon_\alpha = \Sigma(x + \varepsilon_x)\ e^{-(z + \varepsilon_z)} - \Sigma\ x\ e^{-z}. \tag{10}$$

This can be broken into

$$N\ \varepsilon_\alpha = \Sigma\ x\ e^{-(z + \varepsilon_z)} + \Sigma\ \varepsilon_x e^{-(z + \varepsilon_z)} - \Sigma\ x\ e^{-z}. \tag{10a}$$

If we assume that $e^{-\varepsilon_z} \sim 1$, since ε_z will be a small size, we obtain approximately

$$\varepsilon_{\alpha_1} = \Sigma\ \varepsilon_x\ e^{-z}. \tag{10b}$$

It is evident that for $N \to \infty$ the $\epsilon_x \to o$ and $\epsilon_{\alpha_1} \to o$, especially since then $e^{-\epsilon_x} = 1$. We further can assume that ϵ_x will be positive and negative random errors and then ϵ_{α_1} should be small. It may suffice to test ϵ_{α_1} by computations from actual data on hand.

Attention should be further directed towards the fact that α itself is an estimator and the observations x are attached with an error ϵ_x for estimation of $\hat{\alpha}$.

Comparison between the maximum likelihood estimator $\hat{\alpha}$ and the four other estimators has been made for 2 seasons, Summer and Winter, for levels 1 through 24 km and for 4 stations from different climatic regions (Albrook, Canal Zone; Montgomery, Alabama; Berlin, Germany and Thule, Greenland). Then the percentage difference was computed, as the $\hat{\alpha}$ varied in size, especially for the density;

$$\Delta = (\alpha_j - \hat{\alpha})/\hat{\alpha}. \tag{11}$$

where the α_j represents the 4 other estimators introduced previously.

The mean ($\bar{\Delta}$), standard deviation σ_Δ, and the absolute mean difference $|\bar{\Delta}|$ was then calculated for the entire set of estimators. The result is exhibited in Table 1. It is evident that the $\hat{\alpha}_1$ displays the closest approximation of all estimators, as expected, with average deviations less than 3 percent.

Although the minimum variance estimators α^* discloses a similar mean ($\bar{\Delta}$) value, it can be seen that the mean absolute deviation (in percent) is by far higher and the standard deviation each of the 96 estimators (Summer and Winter each) is by far greater than for the α_1. This is no surprise, and illustrates the applicability of the solution by equation (8a) for practical purposes. The computer time to calculate $\hat{\alpha}_1$ was only a minute fraction of the time for $\hat{\alpha}$ and considerably less than for α^*.

It must be further called to attention that although the computation is based upon an order term, the sample sequence need not be known a priori. The method can even be applied to grouped data, reducing merely the number of summations for $\Sigma \, x \, e^{-z}$. It is understood that the x must then be replaced by the upper boundary value, as we are dealing with cumulative distribution in setting $F(z) = F(x)$.

Tables 2 through 4 exhibit the results for α_j and u_j for the station Montgomery as an example that the findings for the individual stations resemble the summary. The survey was also made for the absolute deviation, $\Delta = \alpha_j - \hat{\alpha}$ \hfill (11a)

but the outcome is similar to the one presented and may be omitted.

TABLE 1

Comparison with Maximum Likelihood for α
(1 through 24 km, Summary 4 stations)

		Wind			Density			Temperature		
		$\bar{\Delta}$	σ_Δ	$|\bar{\Delta}|$	$\bar{\Delta}$	σ_Δ	$|\Delta|$	$\bar{\Delta}$	σ_Δ	$|\Delta|$
Summer	Mom.	3.8%	9.2%	8.6%	10.9%	9.8%	13.0%	13.2%	7.9%	14.0
	Min. Var	2.2	11.8	10.7	1.9	9.2	9.2	4.9	13.4	12.2
	Eqn. 8a	-.6	4.3	3.5	-1.3	2.0	2.2	-1.1	2.3	2.4
	Eqn. 9	-5.6	4.1	7.2	-1.4	2.1	2.3	-1.1	2.4	2.4
Winter	Mom.	7.7	6.8	8.9	15.8	9.7	16.3	11.4	9.2	15.1
	Min. Var	-2.7	9.8	9.0	1.0	10.3	8.3	0.0	9.2	8.0
	Eqn. 8a	-2.8	2.6	3.4	-.7	2.8	2.6	-.9	2.6	2.2
	Eqn. 9	-5.6	3.4	6.0	-.6	3.0	2.7	-.2	3.9	3.0

TABLE 2

Comparison with Maximum Likelihood
Wind Speed (1 through 24 km Altitude)

Δ in %

	α			u		
	$\overline{\Delta}$	σ_Δ	$\overline{|\Delta|}$	$\overline{\Delta}$	σ_Δ	$\overline{|\Delta|}$
Summer						
Mom	4.3%	4.4%	5.1%	0.5%	.6%	.0.7%
Min. Var.	13.3	13.4	14.1	1.5	1.6	1.7
Eqn. 8a	-1.3	2.2	2.0	-.2	.3	.3
Eqn. 9	-0.7	1.8	6.7	-1.0	.3	1.0
Winter						
Mom	8.4%	8.6%	9.4%	.6%	.7%	.7%
Min. Var.	-3.5	10.5	9.1	-.0	.9	.7
Eqn. 8a	-1.8	2.7	2.5	-.1	.3	.2
Eqn. 9	-4.4	4.2	5.1	-.6	.8	.6

summer α: Range .13 through .54   u: Range 7.1 through 32.3 m/sec

winter α: Range .10 through .79   u: Range 9.8 through 77.5 m/sec

TABLE 3

Comparison with Maximum Likelihood
Temperature (1 through 24 km altitude)

Δ in %

		α			u		
		$\overline{\Delta}$	σ_Δ	$\overline{|\Delta|}$	$\overline{\Delta}$	σ_Δ	$\overline{|\Delta|}$
Summer	Mom	12.2%	8.1%	12.8%	.03%	.02%	.03%
	Min. Var.	- .8	10.3	7.9	.01	.03	.02
	Eqn.	-1.4	2.1	2.2	-.00	.01	.01
	Eqn.	-1.4	2.1	2.3	.00	.01	.01
Winter	Mom	14.2%	7.1%	14.6%	.06%	.03%	.06%
	Min. Var.	-5.9	6.9	8.3	-.01	.03	.03
	Eqn.	-1.4	1.2	1.6	-.01	.01	.01
	Eqn.	-1.5	1.3	1.7	-.01	.01	.01

summer α: Range .47 through 1.27   u: Range 306.9 through 208.7°K

winter α: Range .33 through .70   u: Range 294.3 through 211.9°K

-436-

TABLE 4

Comparison with Maximum Likelihood
Density (1 through 24 km Altitude)

Δ in %

		α			u	
	$\overline{\Delta}$	σ_Δ	$\overline{|\Delta|}$	$\overline{\Delta}$	σ_Δ	$\overline{|\Delta|}$
Summer Mom	5.6%	11.4%	9.7%	.02%	.04%	.03%
Min. Var.	- .2	8.9	7.0	.01	.04	.03
Eqn.	-1.5	2.8	2.3	-.00	.01	.01
Eqn.	-1.7	2.1	2.4	-.00	.01	.01
Winter Mom	16.8%	10.3%	16.8%	.08%	.05%	.08%
Min. Var.	2.7	13.1	10.5	.02	.07	.06
Eqn.	- .4	3.0	2.4	-.00	.01	.01
Eqn.	- .4	3.1	2.5	-.00	.02	.01

summer α: Range .12 through 2.20   u: Range 1.195 through .059 g/cm³

winter α: Range .043 through 1.80   u: Range 1.309 through .055 g/cm³

4. COMPARISON OF THE OBSERVED AND ANALYTICAL FREQUENCY

The discussion of the extreme values would be incomplete if no mentioning were made of the suitability of the distributions to upper air data. The best method of parameter estimation is to no avail if the distribution model does not satisfy the observations. It was not intended to have a one-model comparison. Hence the Fisher-Tippett II distribution was also employed for judgment. We have

$$F(z) = \exp\left(-z^{-\gamma}\right) \tag{12}$$

with $\qquad z = x/\beta.$ $\qquad\qquad$ (12a)

The moments fit is

$$\left(\frac{s_x}{x_m}\right)^2 + 1 = \Gamma\,(1 - 2/\gamma)\,/\Gamma^2\,(1 - 1/\gamma) \tag{13a}$$

and $\qquad \beta = x_m/\Gamma\,(1 - 1/\gamma).$ $\qquad\qquad$ (13b)

The Maximum likelihood fit renders the equations

$$\ln \beta - \frac{\ln N - \ln\left(\Sigma x^{-\gamma}\right)}{\gamma} = 0 \tag{13c}$$

$$\frac{N}{\gamma} - \Sigma \ln x + \frac{N \Sigma x^{-\gamma} \ln x}{\Sigma x^{-\gamma}} = 0. \tag{13d}$$

It is evident *that* the moments estimators are not simpler to compute than the maximum likelihood estimators as the γ cannot be expressed in explicit form from the gamma function except by approximation. One can therefore calculate the maximum likelihood estimators just as well. The solution must *be* found by iterative procedures.

After determination of the maximum likelihood estimators for the Gumbel distribution (from equation 8a) and for the Fisher-Tippett II curve the analytical F(x) was computed for both systems. The analytical and empirical cumulative frequencies were compared and the deviations checked by the Kolmogorov-Smirnov significance test:

$$D_M = \max \left| F(x) - F(x_e) \right| \tag{14}$$

where $F(x)$ is the analytical and $F(x_e)$ the empirical value for threshold x. It was decided to test on the 95% significant level. Since most of the judgment by the Kolmogorov-Smirnov test for small N is somehow too optimistic, the modification by Lilliefors (1967) has been employed, which leads to significance at some lower maximum D_M. The test criterion D_a was taken from his tables. The deviation was considered significant when $D_M \geq D_a$.

The investigation for Montgomery as the test station disclos ;hat Gumbel's law and Fisher-Tippett II are both good approximations f che wind speed at altitude levels. Only at one altitude level the significance threshold was exceeded. This occurred for every model which was tested, not only one method and is not critical. In a significance test with 48 samples at the 95% level 2 samples may exceed the 95% criterion.

A typical example (Fig. 1) was then selected with the 6 km altitude and winter conditions at Montgomery, Alabama. Although in this case no significant deviation could be noted for the Fisher-Tippett II, the fitting of the observed data appears better for Gumbel's law. The average absolute deviations from the observed data are highest for the Fisher-Tippett II.

The temperature extremes were studied next. This time 3 altitude levels displayed some exceedance of the significance threshold for Gumbel's law. This number may be still considered in line with the selected threshold of significance, especially in comparison with the Fisher-Tippett II law. Only 18 samples out of 48 showed no significant deviation for the latter. Consequently we would conclude that Gumbel's law seems better suited to represent temperature extreme values of upper air data. If the same number of deviations for both methods or the upper air wind speed had occurred one would probably interpret the result as an effect of missing observations, or other data related causes; especially high wind speeds may be missing. This explanation fails for the temperature, however.

A typical example is given in Figure 2 for Montgomery at the 8 km level in summer, when the Fisher-Tippett II comparison provided no significant deviation. Again, the Fisher-Tippett II curve exhibits the largest deviations.

Finally upper air density extremes were considered. This time the number of significant deviations from the observed cummulative frequency was higher for the Gumbel law with eight. This is higher than one would normally expect at this significance level. Had we employed the threshold D_a from the unmodified Kolmogorov test no significance would have turned up.

Figure 1

MONTGOMERY, WIND EXTREMES
ALTITUDE 6 KM, WINTER

Figure 2

MONTGOMERY, TEMPERATURE EXTREMES
ALTITUDE 8KM, SUMMER

Figure 3

MONTGOMERY, DENSITY EXTREMES
ALTITUDE 8 KM, SUMMER
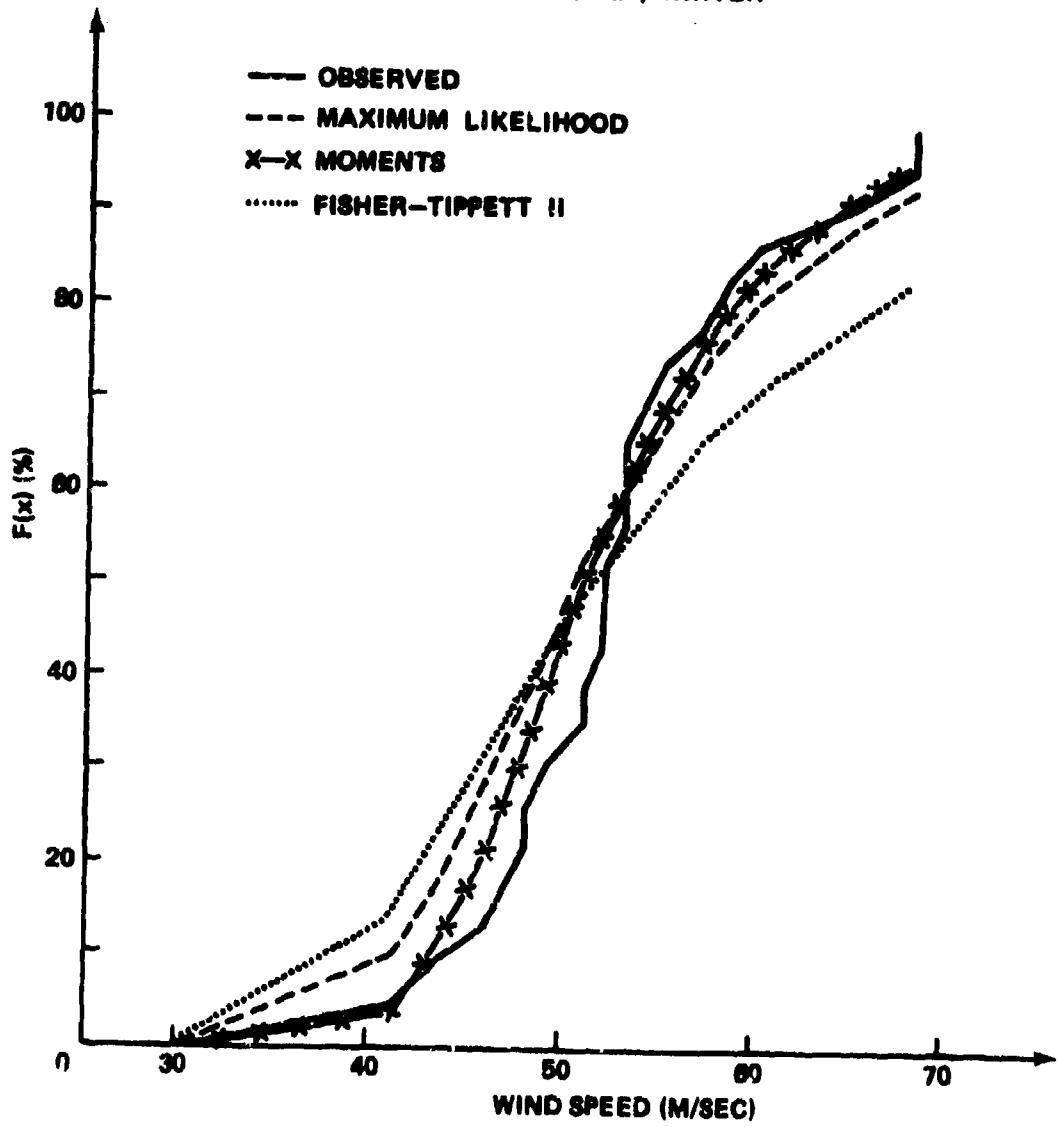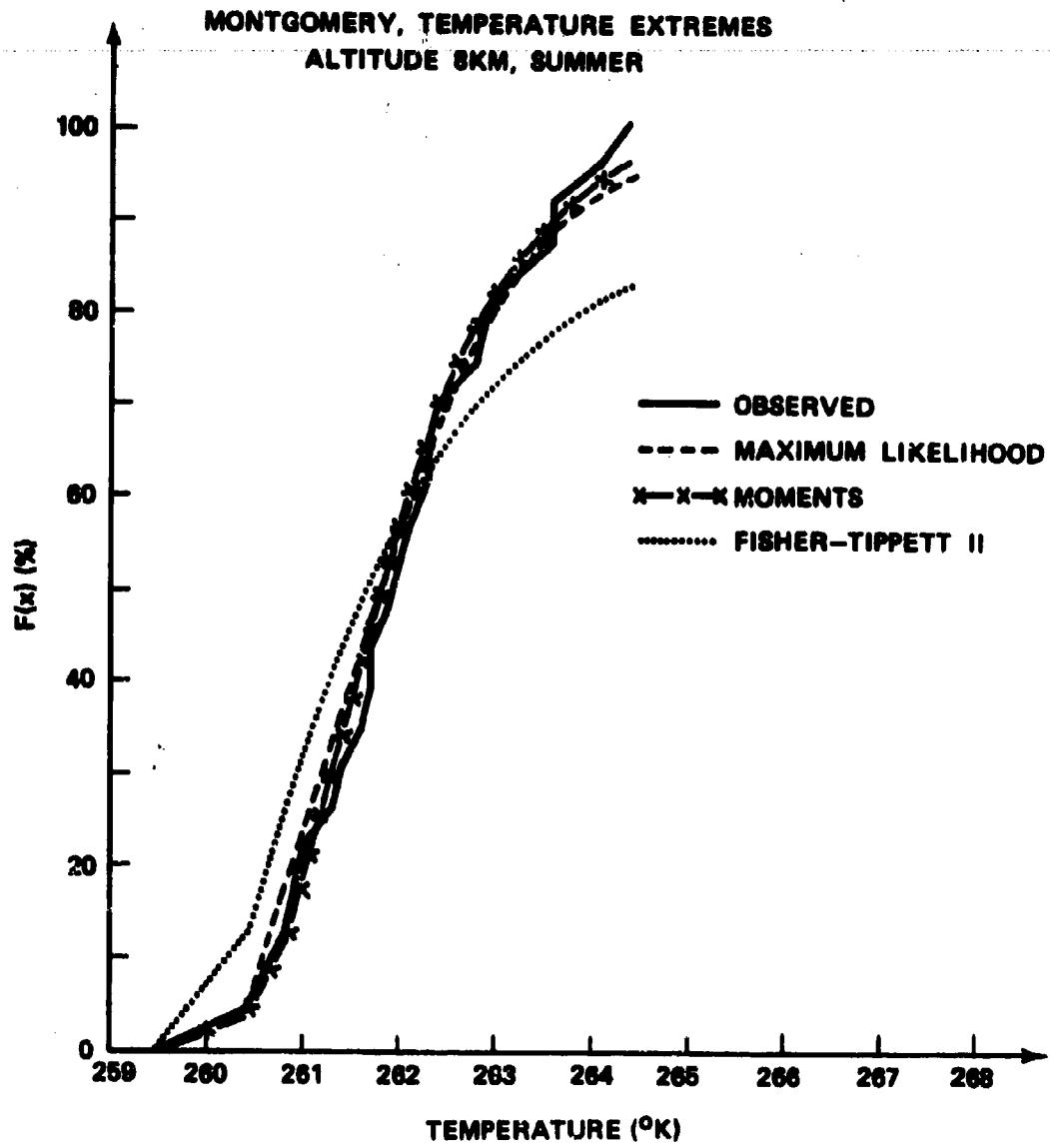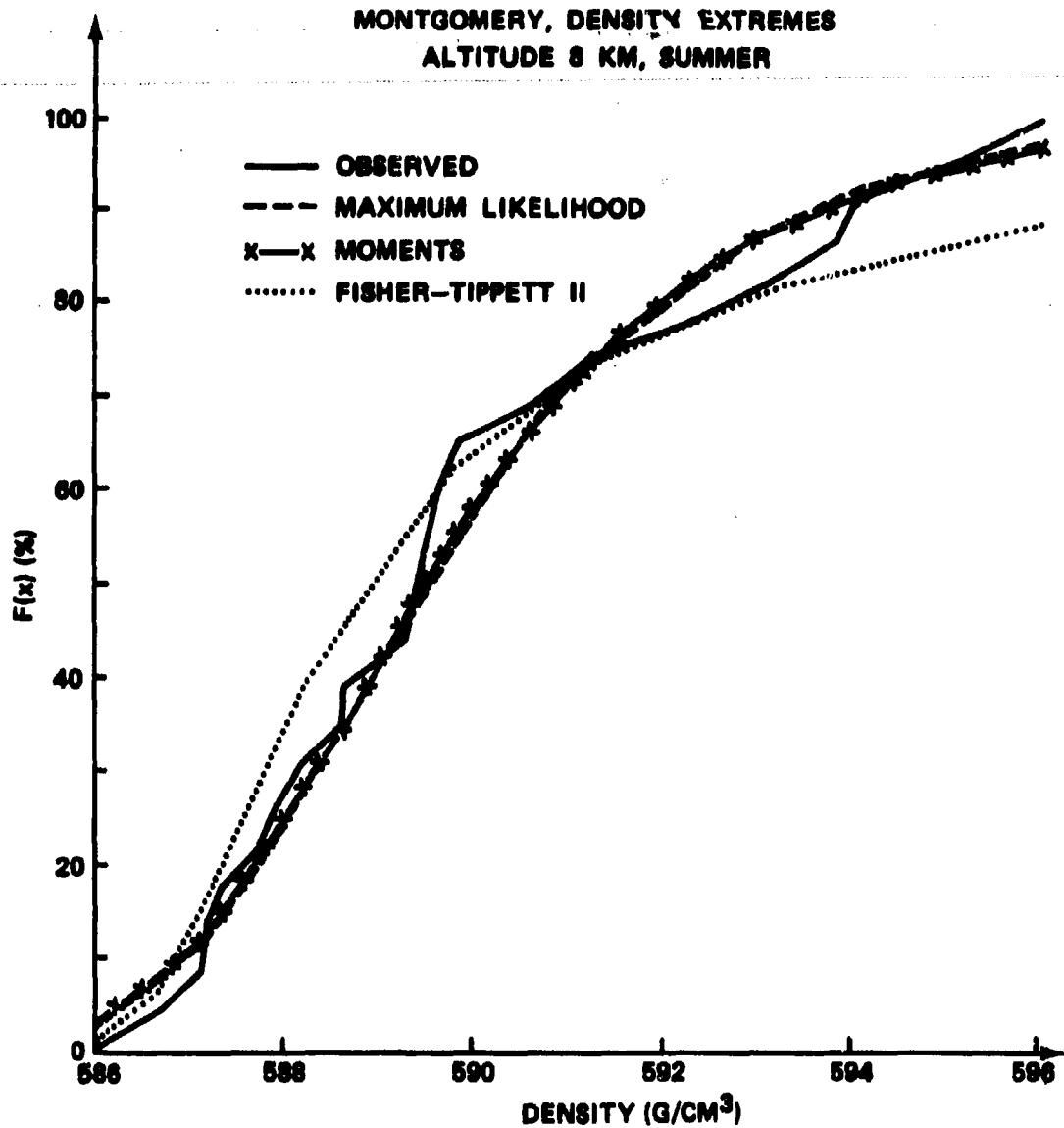
The Fisher-Tippett II behaved similarly to the outcome for the
temperature, only 18 samples showed no significance, although samples
with exceedance for the temperature did not automatically prove significant
for the density. Therefore Gumbel's law seems to fit the observational
data better again.

The higher number of significant deviations for the density may have
some other explanation. The density is an element derived by the gas law
from pressure and temperature. The density extremes may therefore come
from a collective of a bivariate distribution. This modification has
only very recently been studied by Campbell and Tsokos (1972). Their
findings prove that correlated normal maxima are asymptotically uncorrelated.
Thus density extremes would come from uncorrelated pressure and temperature
data even with the existence of correlation between temperature and pressure.
This modification may affect the distribution. Further investigations
seem appropriate before final conclusions can be drawn.

5. CONCLUSIONS: It has been demonstrated in the previous sections
that the maximum likelihood estimators can be approximated by a slight
modification of the likelihood equations employing the analytical
cummulative distribution function. This presents an explicit solution
for the estimator $\hat{a}$ and thus reduces the time of computation considerably.
The method has the advantage that data need not be given in sequence
of data sampling, which is required for Lieblein's minimum variance
linear order statistics estimator. Establishment of an m · k data matrix
can also be waived. Comparison of the two approximate estimators with
the maximum likelihood estimator revealed that especially the first
suggested method gives excellent agreement with the correct solution.
The new estimators have the properties of maximum likelihood estimators.

A final comparison of the analytical cummulative frequency with the
observed one was performed. In this study the analytical curve from the
Fisher-Tippett II extreme value distributions was added. Gumbel's law
and the Fisher-Tippett II model appear equivalent for the wind speed at
the tested significance level. Montgomery data with 2 seasons and 24
altitude levels served as the checking station.

The Gumbel distribution proves significantly better suited for upper
air temperature and density. Even for wind speed the differences between
analytical and observed distributions is smaller for Gumbel's law.
With the availability of an explicit solution for the maximum likelihood
estimators it should not be difficult to apply them to a large number of
samples, as even then costs for electronic data processing stay within
reasonable limits, which would not be true for iteration procedures.

CITED REFERENCES

Campbell, J. W. and Tsokos, Ch. P., 1972 The Asymptotic Distribution of Maxima in Bivariate Samples. Unpubl. Manuscript (Virg. Polytechn. Institute and SU, Black, b., Va.)

Fisher, R. A. and L.H.C. Tippett, 1928 Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample. Proc. Cambridge Phi. Soc., Vol. 24, p. 180-190.

Gumbel, E., 1958 Statistics of Extremes. Columbia University Press, New York, pp. 375.

Kolmogorov, A., 1933 Sulla Determinazione empirica di una legge di distributione. G. Ist. Ital. Attuari, Vol. 4, p. 1-14.

Lieblein, J., 1954 A New Method of Analyzing Extreme-Value Data. NACA-TN-3053, pp. 90

Lilliefors, H. W., 1969 On the Kolmogorov-Smirnov Test for Goodness-of-Fit Journ. Am. Stat. Assoc., Vol. 46, p. 68-78.

Smirnov, N., 1948 Table for Estimating the Goodness-of-Fit of Empirical Distributions. Am. Math. Stat., Vol. 19, p. 279-281.

ACKNOWLEDGMENT

STATISTICAL MODELS FOR H. F. IONOSPHERIC
FORECASTING FOR FIELD ARMY DISTANCES

Richard J. D'Accardi, Robert A. Kulinyi,
U. S. Army Electronics Command
Fort Monmouth, New Jersey
and
Chris P. Tsokos
Department of Statistics and Statistical Laboratory,
Virginia Polytechnic Institute and State University

ABSTRACT. The aim of this paper is to develop statistical models
to forecast short-path oblique incidence (OI) high frequency (HF)
information up to a certain time in advance, by utilizing the observed
vertical incidence data over typical field army distances.

It is shown that there is a strong linear dependence between the
oblique and vertical incidence ionospheric soundings. Linear regression
models have been developed to estimate the oblique incidence soundings from
observed vertical incidence recordings.

It is further shown that ionospheric data of this type is a non-
stationary stochastic realization. A procedure is presented in modeling
such information for the purpose of forecasting one, two, three, ...,
k time slots ahead over a given path. Autoregressive and moving average
forecasting models have been formulated for the 60 km path from Fort
Monmouth to Fort Dix. Confidence bounds have been obtained for both
the linear regression models and the time series analysis of the ionospheric
data.

1. INTRODUCTION. The field army employs many means of communication,
each of which is tailored to fit a particular requirement. Specifically,
H.F. communications provide systems not specifically limited by line-of-sight,
extended distance, or intervening terrain obstacles. However, size and
weight of tactical communications equipment must be kept to a minimum,
and the tactical communicator must contend with a relatively low power
transmitter, and physically small, inefficient low gain antennas.
Ionospheric disturbances, both natural and man-made, further complicate
his problems. The H. F. communicator has, as his only propagation aid,
the monthly predications for undisturbed conditions prepared three
months in advance by the Department of Commerce, and distributed by the
US Army Strategic Communications Command. While valid for long range
planning, they do not account for diurnal variations or disturbed
ionospheric conditions which may harass him. As a result, the U. S.
Army Electronics Command is developing a system to provide tactical
communicators with propagation predictions, in near real-time, and
prepared specifically for the Army area of interest.

The aim of this paper, therefore, is two-fold:

a. to introduce a new statistical concept to the estimation of oblique incidence soundings, either knowing or being able to predict the vertical incidence sounding, and

b. to develop statistical models to forecast either the oblique or vertical incidence soundings over specific paths, or at specific terminals, one, two, three, . . ., k time slots ahead, beginning with a certain origin.

With respect to the first objective, the widely accepted approach to the subject area was to utilize the secant law to estimate the oblique incidence sounding. Simply stated, it is:

Y = X secant φ where

    Y = equivalent oblique incidence frequency
    X = observed vertical incidence frequency
    φ = angle between the oblique ray path and the normal to the
        ionosphere at the path mid-point

This is Snell's law used under the assumptions of constant ionospheric layer height, no collisions, no magnetic field effects, and a spectral reflection at the path mid-point. If one assumes a curved ionosphere (for longer paths), however, a modification to the above equation, depending upon the electron density profile, is of the form:

Y = k X secant η where a practical range of k is:

    $1.00 \leq k \leq 1.30$

In view of the poor results obtained by Krause, et al, [7], using secant φ, the probable need for mid-point data, and the dependence on electron profile density, a more practical approach for relating VI and OI data was developed, using regression techniques. Functional relationships were derived for each experiment as a prelude to the forecasting problem.

In Section 3, we shall give the regression models for the overall mean and the 6th day measurements including reciprocal path data for the 60 Km, 200 Km, and 500 Km paths. The confidence intervals of these linear models are also given.

We have shown that there is a strong linear dependence between oblique and vertical incidence soundings at all the paths investigated. The regression models developed show that one can accurately estimate the equivalent oblique incidence sounding for a given value of the vertical incidence data. What remains is the need for a model that, under certain realistic

-446-

conditions, will enable one to forecast the vertical incidence ionospheric soundings and, from the forecast, to estimate the corresponding OI sounding. This is the aim of the second part of our paper.

It is shown that for the 60 Km experiment, both the oblique and vertical incidence recordings are non-stationary stochastic realizations. That is, they form a discrete time series which is not in statistical equilibrium. We propose a procedure to handle this type of information and to investigate the possibility of characterizing our data with an autoregressive process, a moving average model, or a mixture of autoregressive-moving average processes.

In Section 4, we present a systematic presentation of analyzing ionospheric soundings for the purpose of forecasting. An autoregressive model has been developed for the discrete realization representing the 6th day observed oblique incidence critical frequencies for the 60 Km path, Fort Monmouth, N. J. - Fort Dix, N. J., in Section 5. The complete procedure of fitting such a model is given, along with its confidence intervals. In Section 6, we develop a moving average model that characterizes the behavior of the overall vertical incidence soundings for the 60 Km experiment. Forecasting models for the 200 Km and 500 Km paths from Fort Monmouth, N. J., to Aberdeen Proving Ground, Md., and from Fort Monmouth, N. J., to Camp Drum, N. Y., respectively, are presently being developed.

The remainder of this paper was reproduced photographically from the manuscript submitted by the author.
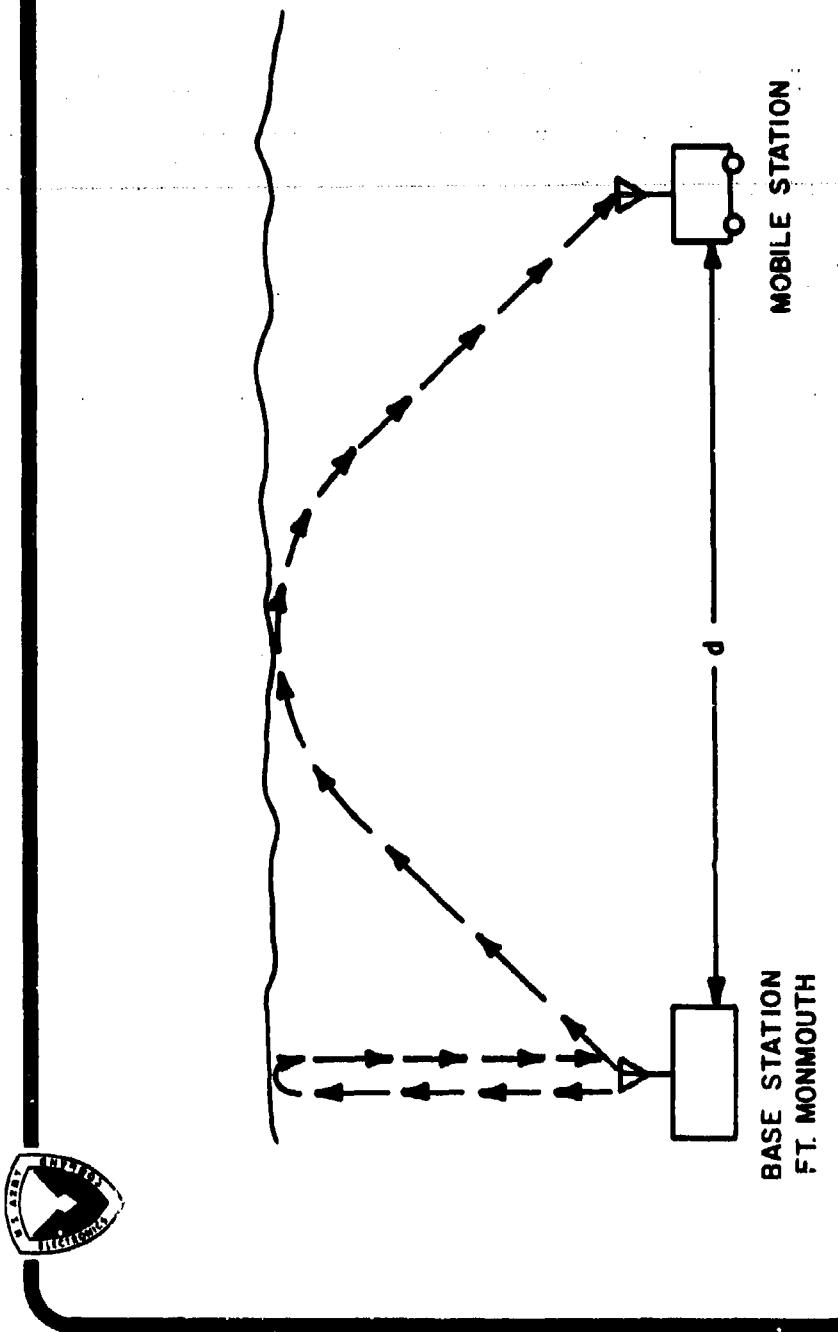
2. DESIGN OF THE EXPERIMENT.

To accomplish the task of developing both a functional relationship between OI and VI maximum observed frequencies (MOF) and a forecasting model, the Communications/ADP Laboratory, of the U. S. Army Electronics Command, has been involved in an extensive collection of VI and short-path OI ionospheric data at three different distances with Fort Monmouth, N. J., as the base station. Experimentation was performed in the 2-16 MHz range, using two ionosondes, one as a fixed terminal and the other as a mobile terminal, as shown in Figure 2.1. The mobile terminal was situated at Fort Dix, N. J., establishing a 60 Km path; at Aberdeen Proving Ground, Md., to establish a 200 Km path; at Camp Drum, N. Y., to establish a nominal 500 Km path (Figure 2.2). These particular distances were chosen to fall within the idealized 300 x 300 kilometers tactical Field Army area of responsibility. The nominal 500 Km path (actually 440 Km) is representative of the diagonal or largest internal communications path within the area of responsibility.

Each terminal made scheduled soundings every ten minutes for 9, 19, and 18-day experiments, respectively. 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 (Figures 2.1 and 2.2). Both ionosondes were synchronized to the WWV (HF), (National Bureau of Standards) time standard so that the "remote" sounder scans would be precise with the Fort Monmouth terminal. The number of days each experiment was performed has no significance with respect to the results obtained, but was a matter of funding. The basic instruments used were two Granger Associates Model 3905-5 Ionospheric Sounders, matched with wide response delta antennas.
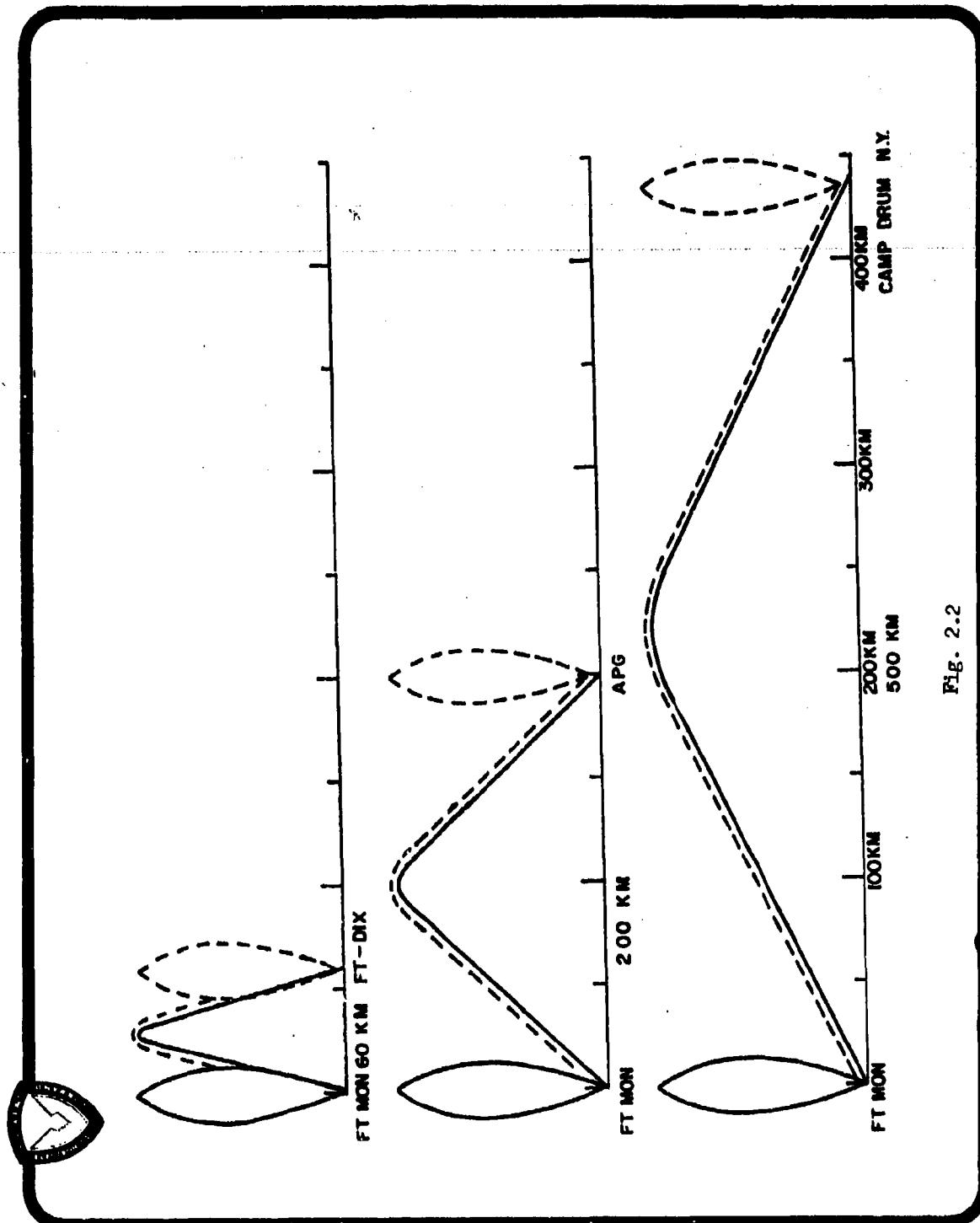
The frequency range of the ionosondes was limited from 2-16 MHz, in three octaves, with 400 discrete frequency channels per octave. Transmissions consisted of successively "stepping" through the channels of each octave with a pulse width of 100 micro-seconds to maximize the system sensitivity. The data is a recording of the time delay from ionosonde to ionospheric reflecting layer and return. Time delay is a measure of the virtual height of reflection from the layer. The trace of the returned pulse on a scale of frequency versus time delay (virtual height) is the ionogram record, Figure 2.3. Iono-gram records of the data were taken on 35mm film at Fort Monmouth and on light sensitive oscillograph paper at the remote terminals. After collection and development, the ionograms were scaled for the extraordinary critical frequencies, $f_x F_2$, as shown in Figure 2.3. The $f_x F_2$ data was then compiled for computer analysis and for comparison between the observed VI and observed OI critical frequencies.

The experiment results were dependent upon ionospheric conditions and man-made noise. Conditions were characterized by the Space Disturbance Forecast Center, ITS, Boulder, Colorado, as generally undisturbed, but some interference occurred. Some data (ionograms) were unreadable due to man-made noise, solar and geomagnetic activity. For those few records which were unreadable (though signal was detected), simulated data was prepared. The occurrence of obscured data was negligible over the experiments.

FIELD TESTS FOR A NEAR-REAL TIME
IONOSPHERIC FORCASTING SCHEME

Fig. 2.1

Fig. 2.2

$f_x f_2$
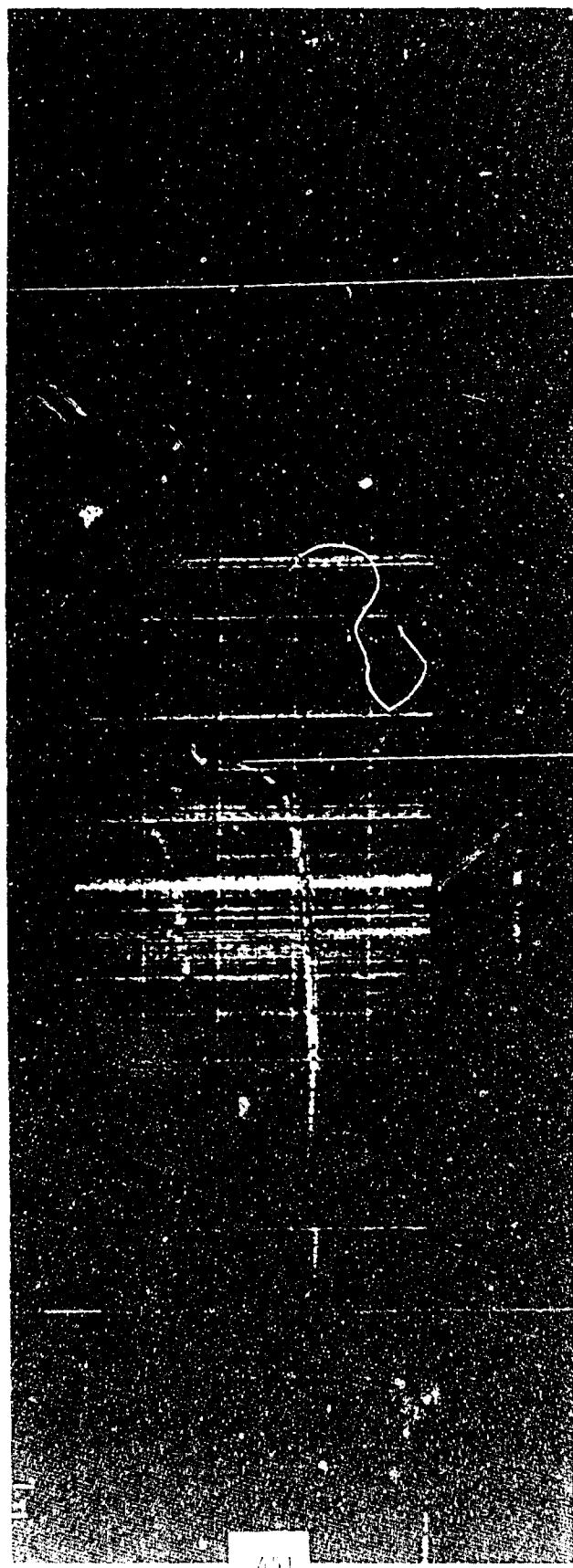
Fig. 2.3

3. REGRESSION MODEL

It has been previously argued that the secant law approach for estimating short-oblique incidence HF information by utilizing observed vertical incidence data does not take into consideration the stochastic behavior of the problem. That is, the secant φ approach is a deterministic characterization of the experiment. The complexity of such experiments is subject to human error, ionospheric changes, assumptions of constant layer height, of no collisions, and of no magnetic effects, among others. It is quite unrealistic, therefore, to consider such a phenomenon from a deterministic point of view.

It has been shown (see Table 3.1 that there exists a strong linear dependence between the oblique incidence and vertical incidence soundings for all the distances investigated in the experiment, that is, 60 Km, 200 Km, and 500 Km.

Thus, in view of the strong linear dependence shown by the oblique and vertical incidence ionospheric soundings, we formulated pairs of reciprocal linear regression models for the overall averages and for the 6th day measurements for the three paths. These models are shown in Table 3.2, where X and Y represent the vertical and oblique incidence ionospheric data.

The usefulness of such models is quite clear. That is, at a given distance, if one is willing to assume similar ionospheric conditions, seasonal variations, and terrain, a very good estimate of the oblique incidence sounding can be obtained by knowing the vertical incidence sounding.

Confidence limits were obtained for each of the models shown in Table 3.2. More specifically, 95% confidence intervals have been computed and plotted for the overall averages and the 6th day measurements for each of the paths studied. Figures 3.3 through 3.14 show the formulated regression models with their computed confidence limits.

For a complete study of the regression analysis modeling of ionospheric data, refer to D'Accardi and Tsokos [6].

Correlation Between Vertical Incidence And
Oblique Incidence Ionospheric Data

Experiment Day	Nominal 60 Km Path		Nominal 200 Km Path		Nominal 500 Km Path	
	Ft. Monmouth to Ft. Dix	Ft. Dix to Ft. Monmouth	Ft. Monmouth to Aberdeen P.G.	Aberdeen P.G. to Ft. Monmouth	Ft. Monmouth to Camp Drum	Camp Drum to Ft. Monmouth
1	.993	.996	.957	.973	.995	.975
2	.989	.993	.985	.985	.994	.988
3	.990	.995	.989	.987	.993	.991
4	.989	.996	.994	.988	.993	.985
5	.995	.998	.989	.990	.989	.991
6	.990	.998	.989	.965	.990	.935
7	.992	.997	.988	.987	.995	.993
8	.989	.998	.991	.989	.994	.959
9	.996	.997	.991	.987	.979	.938
10			.974	.987	.993	.990
11			.983	.992	.995	.985
12			.974	.970	.986	.992
13			.992	.989	.995	.967
14			.991	.973	.978	.985
15			.957	.967	.977	.989
16			.986	.980	.987	.835
17			.986	.932	.895	.932
18			.978	.978	.992	
19			.974	.850		
Overall Path Data	.999	.999	.999	.998	.988	.999

TABLE 3.1

453-

Regression Models Relating Observed Vertical Incidence,
(X_{ij}), And Equivalent Oblique Incidence,(Y_{ij}), Ionospheric
Data

Nominal Distance	Path	Overall Averages Model	6th Day Model
60 Km	Ft. Monmouth, N.J. to Ft. Dix, N.J.	$Y_{oj} = 8.148 +1.001\ (X_{oj}-8.161)$ $j = 1,2,\ldots., 85$	$Y_{6j} = 8.752 +1.019\ (X_{5j}-8.151)$ $j = 1,2,\ldots.,65$
50 Km	Ft. Dix, N.J. to Ft. Monmouth, N.J.	$Y_{oj} = 8.877 +1.020\ (X_{oj}-8.856)$ $j = 1,2,\ldots.,85$	$Y_{6j} = 9.458 +1.023\ (X_{6j}-9.420)$ $j = 1,2,\ldots.,85$
200 Km	Ft. Monmouth, N.J. to Aberdeen P.G.	$Y_{oj} = 7.461 +.995\ (X_{oj}-7.442)$ $j = 1,2,\ldots.,144$	$Y_{6j} = 7.609 +.921\ (X_{6j}-7.653)$ $j = 1,2,\ldots.,144$
200 Km	Aberdeen P.G. to Ft. Monmouth, N.J.	$Y_{oj} = 7.516 +1.020\ (X_{oj}-7.467)$ $j = 1,2,\ldots.,144$	$Y_{6j} = 7.686 -1.066\ (X_{6j}-7.512)$ $j = 1,2,\ldots.,144$
500 Km	Ft. Monmouth, N.J. to Camp Drum, N.Y.	$Y_{oj} = 7.528 +1.160\ (X_{oj}-7.199)$ $j = 1,2,\ldots.,144$	$Y_{6j} = 7.768 +1.197\ (X_{6j}-7.419)$ $j = 1,2,\ldots.,141$
500 Km	Camp Drum, N.Y. to Ft. Monmouth, N.J.	$Y_{oj} = 7.525 +1.058\ (X_{oj}-6.618)$ $j = 1,2,\ldots.,144$	$Y_{6j} = 6.273 +1.043\ (X_{6j}-5.650)$ $j = 1,2,\ldots.,144$

TABLE 3.2

LINEAR REGRESSION MODEL FOR 60 Km PATH WITH 95% CONFIDENCE INTERVALS — FT. MONMOUTH TO FT. DIX FOR THE 6TH DAY.

$\hat{Y} = 8.752 + 1.019(X - 8.771)$

$\hat{Y}$ = ESTIMATED OBLIQUE INCIDENCE DATA

X = OBSERVED VERTICAL INCIDENCE DATA

OBSERVED VERTICAL INCIDENCE DATA — MHz

ESTIMATED OBLIQUE INCIDENCE DATA — MHz

Fig. 3.3

$\hat{Y} = 9.458 + 1.023(X - 9.420)$

$\hat{Y}$ = ESTIMATED OBLIQUE
INCIDENCE DATA

X = OBSERVED VERTICAL
INCIDENCE DATA

LINEAR REGRESSION MODEL FOR 60 Km PATH WITH 95%
CONFIDENCE INTERVALS – FT. DIX TO FT. MONMOUTH
FOR THE 6TH DAY

OBSERVED VERTICAL INCIDENCE DATA – MHz

ESTIMATED OBLIQUE INCIDENCE DATA – MHz

Fig. 3.4

$\hat{Y} = 8.148 + 1.001(X - 8.161)$

$\hat{Y}$ = ESTIMATED OBLIQUE INCIDENCE DATA

X = OBSERVED VERTICAL INCIDENCE DATA

LINEAR REGRESSION MODEL FOR 60 Km PATH WITH 95% CONFIDENCE INTERVALS –FORT MONMOUTH – FORT DIX OVERALL AVERAGE FOR 9 DAYS.

OBSERVED VERTICAL INCIDENCE DATA–MHz

ESTIMATED OBLIQUE INCIDENCE DATA–MHz

Fig. 3.5

Fig. 3.6

Fig. 3.7

$\hat{Y} = 7.686 + 1.066 (X - 7.512)$

$\hat{Y} =$ ESTIMATED OBLIQUE INCIDENCE DATA

X = OBSERVED VERTICAL INCIDENCE DATA

LINEAR REGRESSION MODEL FOR 200 Km PATH WITH 95% CONFIDENCE INTERVALS - APG TO FT. MONMOUTH FOR THE 6TH DAY

OBSERVED VERTICAL INCIDENCE DATA — MHz

ESTIMATED OBLIQUE INCIDENCE DATA—MHz

Fig. 3.8

LINEAR REGRESSION MODEL FOR 200 Km PATH WITH 95% CONFIDENCE INTERVALS – FT. MONMOUTH TO APG OVERALL AVERAGE FOR 19 DAYS

$\hat{Y} = 7.461 + 0.995\ (X-7.442)$

$\hat{Y}$ = ESTIMATED OBLIQUE INCIDENCE DATA

X = OBSERVED VERTICAL INCIDENCE DATA

Fig. 3.9

LINEAR REGRESSION MODEL FOR 200 Km PATH WITH 95%
CONFIDENCE INTERVALS - APG TO FT. MONMOUTH
OVERALL AVERAGE FOR 19 DAYS

Fig. 3.10

$\hat{Y} = 7.516 + 1.020(X - 7.407)$

$\hat{Y}$ = ESTIMATED OBLIQUE
        INCIDENCE DATA

X = OBSERVED VERTICAL
    INCIDENCE DATA

OBSERVED VERTICAL INCIDENCE DATA - MHz

ESTIMATED OBLIQUE INCIDENCE DATA - MHz

Fig. 3.11

LINEAR REGRESSION MODEL FOR 500 Km PATH AND 95%
CONFIDENCE LIMITS FOR:
CAMP DRUM, N.Y. – FORT MONMOUTH, N.J.
FOR THE 6TH DAY

$\hat{Y} = 6.273 + 1.043\ (X - 5.650)$

$\hat{Y}$ = ESTIMATED OBLIQUE
INCIDENCE DATA

X = OBSERVED VERTICAL
INCIDENCE DATA

OBSERVED VERTICAL INCIDENCE DATA – MHz

ESTIMATED OBLIQUE INCIDENCE DATA–MHz

Fig. 3.12

$\hat{Y} = 7.528 + 1.160 (X - 7.199)$

$\hat{Y} =$ ESTIMATED OBLIQUE
INCIDENCE DATA

$X =$ OBSERVED VERTICAL
INCIDENCE DATA

LINEAR REGRESSION MODEL FOR 500 Km PATH AND 95 %
CONFIDENCE LIMITS FOR:
FORT MONMOUTH, N.J. – CAMP DRUM, N.Y.
OVERALL AVERAGE FOR 18 DAYS

OBSERVED VERTICAL INCIDENCE DATA – MHz

ESTIMATED OBLIQUE INCIDENCE DATA – MHz

Fig. 3.13

$\hat{Y} = 7.525 + 1.058 (X - 6.618)$

$\hat{Y}$ = ESTIMATED OBLIQUE INCIDENCE DATA

X = OBSERVED VERTICAL INCIDENCE DATA

LINEAR REGRESSION MODEL FOR 500 Km PATH AND 95% CONFIDENCE LIMITS FOR: CAMP DRUM, N.Y. – FORT MONMOUTH, N.J. OVERALL AVERAGE FOR 15 DAYS

OBSERVED VERTICAL INCIDENCE DATA – MHz

ESTIMATED OBLIQUE INCIDENCE DATA – MHz

Fig. 3.14

4. TIME SERIES MODELING: IDENTIFYING THE STOCHASTIC REALIZATION

In a given physical situation such as in the ionospheric sounder project, we have available a time series, say, $x_1, x_2, \ldots, x_n$, of n observations. Our aim is to obtain a suitable difference equation or model that will accurately represent the true underlying process which generated the ionospheric soundings, x_t, $t=1,2,\ldots,n$. First, we must identify whether the series x_t exhibits stationary or non-stationary properties. That is, when we speak of stationary time series, we imply that the statistical properties of the series are independent of absolute time. A graphical representation of the ionospheric soundings would be of some aid in exercising judgment about the behavior of the data. Of greater importance is the sample autocorrelation function given by:

$$r_{xx}(k) = \frac{c_{xx}(k)}{c_{xx}(0)}, \quad k = 0,1,2,\ldots,n-1$$

where $c_{xx}(k)$ is the sample autocovariance function defined by:

$$c_{xx}(k) = \frac{1}{n} \sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t+k} - \bar{x}), \quad k=0,1,2,\ldots n-1,$$

of the observed soundings. If the ionospheric soundings are stationary, the sample autocorrelation function would exhibit fairly rapid dampening. Furthermore, one can apply various statistical tests to check for nonstationary properties. We have used Kendall's tau test [5].

If the ionospheric soundings were not in statistical equilibrium (stationary), then we can filter out the non-stationary components by using various difference filters. In all cases, that is, the 60 km, 200 km, and 500 km experiments, the resulting data was shown to contain non-stationary components.

A general difference filter is given by:

$$y_t = (1-B)^d x_t,$$

where B is a shift operator and d is the order of the filter. When $d=0$, this will indicate that the ionospheric data is stationary; $d=1$, will indicate that a first difference filter is necessary to filter the original series, and so on. For the ionospheric soundings information, we used filters up to $d = 2$.

The procedure to determine the proper value for d is to compute the first differences of the original data, x_t, $t=1,2,\ldots,n$. That is, we process x_t through a first difference filter:

$$y_t = (1-B) x_t = x_t - x_{t-1},$$

which will have (n-1) observations, and then through a second difference filter:

$$w_t = (1-B)^2 x_t = x_t - 2x_{t-1} + x_{t-2}$$

which will have (n-2) observations. For the original soundings, x_t, and the filtered y_t and w_t, we calculate the sample autocorrelation function and

Kendall's tau [5]. By observing the sample autocorrelation function of the
original series, the filtered soundings, and the results of the trend test, we
can infer a suitable value for d, that is, the degree of "differencing"
necessary to induce the sample autocorrelation function to dampen out fairly
rapidly and cause Kendall's tau test not to be significant.

For predicting or forecasting oblique incidence data for one or more time
slots in advance, the initial step is to determine the particular process
that characterizes our data. There are three basic models that are candi-
dates for this purpose:

   a.   The Autoregressive Process
   b.   The Moving Average Process
   c.   The Mixed Autoregressive-Moving Average Process.

A discrete m-order autoregressive model is of the form:

$$x_t - \mu = \alpha_1(x_{t-1} - \mu) + \alpha_2(x_{t-2} - \mu) + \ldots + \alpha_m(x_{t-m} - \mu) + Z_t \qquad (4.1)$$

where x_t is the autoregressive series which is being generated by the series
Z_t, a purely random process, α_1, α_2, $\ldots$, α_m are the parameters of the non-
ordered process, and μ is the expected value of the series. Such a process
assumes that the current value x_t of the soundings has resulted from a linear
sum of past values of the series. Such a process assumes that the current
value, x_t, of the soundings has resulted from a linear sum of past values of
the series, together with an independent error term, Z_t, not connected with
the past.
A discrete q-order moving average process is given by:

$$x_t - \mu = Z_t - \beta_1 Z_{t-1} - \beta_2 Z_{t-2} - \ldots - \beta_q Z_{t-q} . \qquad (4.2)$$

This process is a weighted sum of a random series, Z_t. Each realization
(oblique incidence sounding, x_t) is made linearly dependent on a Z_t and on
one or more previous Z's. Also, μ is the expected value of x_t, and β_1, β_2,
$\ldots$, β_q are the parameters of the model.

The mixed model consists of the autoregressive and moving average model
where m is independent of q.

We shall discuss in some detail a procedural approach in fitting an
autoregressive model to the ionospheric data series. A similar approach can
be followed to formulate the procedure for the moving average and mixed
process with minor changes.

4.1   The Autoregressive Process

The autoregressive model previously defined can be adapted to
represent the characterization of ionospheric data for the purpose of fore-
casting. Discussing the theory of the general m^{th} order model is quite
complicated and, therefore, we shall first give a brief discussion of the
second order model which is quite useful in many physical situations.

The second order discrete autoregressive process may be written as:

$$x_t - \mu = \alpha_1(x_{t-1} - \mu) + \alpha_2(x_{t-2} - \mu) + Z_t . \qquad (4.3)$$

If, at the initial stage, it was necessary to filter the ionospheric data to have the information in statistical equilibrium, then we must place certain restrictions on estimating the parameters of the model to make sure that our series remains stationary. To obtain these restrictions on the parameters, we use the concept of $\mathbf{Z}$-transforms, [4], to obtain the characteristic equation of the process. Solving the characteristic equation, we can place conditions on its roots so that the fitted model will not violate the assumption of stationarity.

The $\mathbf{Z}$-transform of equation (4.3) is given by:

$$(1-\alpha_1\mathbf{Z}^{-1}-\alpha_2\mathbf{Z}^{-2}) \, (x_t-\mu) = Z_t \, ,$$

and its transfer function $H(\mathbf{Z}^{-1})$ is given by:

$$H(\mathbf{Z}^{-1}) = \frac{1}{(1-\alpha_1\mathbf{Z}^{-1}-\alpha_2\mathbf{Z}^{-2})}$$

Thus, the characteristic equation of the second-order autoregressive model is:

$$\psi^2 - \alpha_1\psi - \alpha_2 = 0$$

whose roots are given by:

$$\zeta_1 = \frac{\alpha_1 + \sqrt{\alpha_1^2+4\alpha_2}}{2} \quad \text{and} \quad \zeta_2 = \frac{\alpha_1 - \sqrt{\alpha_1^2+4\alpha_2}}{2} \, .$$

In order for the second-order model to be stationary, we must restrict the estimates of the parameters α_1 and α_2 so that the roots of equation (4.4) will be contained within a unit circle, that is, $|\zeta_1|$ and $|\zeta_2|$ must be less than one. This is equivalent to having α_1 and α_2 lie in a triangular region formed by $\alpha_1+\alpha_2 < 1$, $\alpha_2-\alpha_1 < 1$, and $-1 < \alpha_2 < 1$. For additional details see [1], [2], and [3].

A similar approach can be carried out by considering models of higher order. The $\mathbf{Z}$-transform of the mth order autoregressive (4.1) process is given by:

$$(1-\alpha_1\mathbf{Z}^{-1}-\alpha_2\mathbf{Z}^{-2}- \ldots - \alpha_m\mathbf{Z}^{-m}) \, (x_t-\mu) = Z_t \, . \tag{4.5}$$

The transfer function of (4.5) is of the form:

$$H(\mathbf{Z}^{-1}) = \frac{1}{(1-\alpha_1\mathbf{Z}^{-1}-\alpha_2\mathbf{Z}^{-2}- \ldots - \alpha_m\mathbf{Z}^{-m})} \, . \tag{4.6}$$

The characteristic equation of the mth order process is given by:

$$\psi^m -\alpha_1\psi^{m-1} -\alpha_2\psi^{m-2} - \ldots - \alpha_m = 0 \, . \tag{4.7}$$

Thus, for the general finite autoregressive model to be in statistical equilibrium, we must estimate the parameters of the process so that the roots of equation (4.7) must lie within a unit circle.

4.2 The Fitting Procedure

The initial stage in developing any one of the three models under consideration usually involves deciding the order, m, of the model, and then, given m, estimating the parameters, μ, α_1, α_2, ..., α_m.

The criterion for selecting the best order which characterizes the given series is based upon the residual variance. We proceed by estimating the parameters of the model for different orders, and then the residual variances are computed and plotted against the order of the process. The minimum residual variance will correspond to the order of the model which best describes the ionospheric soundings. Thus, for the autoregressive process, it is necessary to first study the estimation of the parameters of the model.

To estimate the parameters of this process, we can use the method of maximum likelihood. We assume that the Z_t process is normal. Then, for a fixed m, the joint probability density function of the variates, Z_{m+1}, Z_{m+2}, ..., Z_n is given by:

$$f_{m+1,\ldots,n}(Z_{m+1}, Z_{m+2}, \ldots, Z_n) = \frac{1}{(\sqrt{2\pi}\,\sigma_z)^{\frac{n-m}{2}}}\, e^{-\frac{1}{2\sigma_z^2} \sum_{t=m+1}^{n} Z_t^2} \quad ,$$

when the expected value of Z_t is zero and its variance is σ_z^2. Changing from the Z variables to the X variables, according to equation (4.1), we have:

$$f_{m+1,\ldots,n}(x_{m+1}, x_{m+2}, \ldots, x_n \mid x_1, x_2, \ldots, x_m)$$
$$= \frac{1}{(\sqrt{2\pi}\,\sigma_z)^{\frac{n-m}{2}}}\, e^{-\frac{1}{2\sigma_z^2} \sum_{t=m+1}^{n} \left\{ (x_t-\mu) - \alpha_1(x_{t-1}-\mu) - \ldots - \alpha_m(x_{t-m}-\mu) \right\}^2} \quad ,$$

$$(4.8)$$

which is the joint probability density function of $x_{m-1}, \ldots, x_n$ conditional on $X_1 = x_1$, $X_2 = x_2$, ..., $X_m = x_m$. Thus, to obtain the joint probability density function of X_1, X_2, ..., X_n, it is only necessary to multiply equation (4.8) with the density of X_1, X_2, ..., X_m. Since m, for practical applications, is usually small, the net effect of not carrying out this multiplication is small and will be omitted. For details, see [3].

The Log-likelihood function of the process may be written as follows:

$$L(\mu, \alpha_1, \alpha_2, \ldots, \alpha_m \mid x_1, x_2, \ldots, x_n) = -(n-m)\ell n \sqrt{2\pi} - (n-m)\ell n\, \sigma_z$$

$$- \frac{1}{2\sigma_z^2} \sum_{t=m+1}^{n} \left\{ (x_t-\mu) - \alpha_1(x_{t-1}-\mu) - \ldots - \alpha_m(x_{t-m}-\mu) \right\}^2 \quad . \tag{4.9}$$

The sum-of-squares function, given by:

$$S(\mu, \alpha_1, \alpha_2, \ldots, \alpha_m \mid x_1, x_2, \ldots, x_n)$$

$$= \sum_{t=m+1}^{n} \left\{ (x_t - \mu) - \alpha_1(x_{t-1} - \mu) - \ldots - \alpha_m(x_{t-m} - \mu) \right\}^2, \tag{4.10}$$

is needed to estimate the parameters of the model. Differentiating equation (4.10), with respect to μ, α_1, α_2, ..., α_m, setting them equal to zero, and solving the m+1 system of equations, we can obtain their maximum likelihood estimates.

For the second-order autoregressive process, we differentiate equation (4.10) with respect to μ, α_1, and α_2, and obtain the following normal equations:

a. $(\bar{x}_3 - \hat{\mu}) = \hat{\alpha}_1(\bar{x}_2 - \hat{\mu}) + \hat{\alpha}_2(\bar{x}_1 - \hat{\mu})$,

b. $\sum\limits_{t=3}^{n}(x_t - \hat{\mu})(x_{t-1} - \hat{\mu}) = \hat{\alpha}_1 \sum\limits_{t=3}^{n}(x_{t-1} - \hat{\mu})^2 + \hat{\alpha}_2 \sum\limits_{t=3}^{n}(x_{t-1} - \hat{\mu})(x_{t-2} - \hat{\mu})$,

and

c. $\sum\limits_{t=3}^{n}(x_t - \hat{\mu})(x_{t-2} - \hat{\mu}) = \hat{\alpha}_1 \sum\limits_{t=3}^{n}(x_{t-1} - \hat{\mu})(x_{t-2} - \hat{\mu}) + \hat{\alpha}_2 \sum\limits_{t=3}^{n}(x_{t-2} - \hat{\mu})^2$,

where

$$\bar{x}_j = \frac{1}{n-3} \sum_{t=3}^{n} x_{t-3+j} , \quad j=1,2,3 .$$

Since $\bar{x}_1$, $\bar{x}_2$, and $\bar{x}_3$ are usually close to the overall mean, $\bar{x}$, we can use it as an approximate estimate of μ. Furthermore, we can obtain good approximate estimates of (a) through (c) using the sample autocorrelation function at lag one, $r_{xx}(1)$. That is,

$$\left. \begin{aligned} c_{xx}(1) &\approx \hat{\alpha}_1 c_{xx}(0) + \hat{\alpha}_2 c_{xx}(1) \ , \\ \text{and} \quad c_{xx}(2) &\approx \hat{\alpha}_1 c_{xx}(1) + \hat{\alpha}_2 c_{xx}(0) \ . \end{aligned} \right\} \tag{4.11}$$

The autocovariance is an even function, thus, we can write equation (4.11) as follows:

$$c_{xx}(j) \approx \hat{\alpha}_1 c_{xx}(j-1) + \hat{\alpha}_2 c_{xx}(j-2) \ , \quad j = 1,2 . \tag{4.12}$$

An approximate estimate of the parameters α_1 and α_2 is given by:

$$\left. \begin{aligned} \hat{\alpha}_1 &\approx \frac{r_{xx}(1)[1 - r_{xx}(2)]}{1 - r_{xx}^2(1)} \ , \\ \text{and} \quad \hat{\alpha}_2 &\approx \frac{r_{xx}(2) - r_{xx}^2(1)}{1 - r_{xx}^2(1)} \ . \end{aligned} \right\} \tag{4.13}$$

Also, an estimate of the residual sum of squares can be obtained in terms of the sample autocovariance function. That is:

$$S(\hat{\mu}, \hat{\alpha}_1, \hat{\alpha}_2) \approx (n-2) \left\{ c_{xx}(0) - \hat{\alpha}_1 c_{xx}(1) - \hat{\alpha}_2 c_{xx}(2) \right\} , \tag{4.14}$$

and the residual variance of Z_t is given by:

$$S_z^2 = \frac{1}{n-5} \; S(\hat{\mu}, \hat{\alpha}_1, \hat{\alpha}_2).$$

Similar expressions can be obtained for estimating the parameters for the general finite autoregressive model. The normal equations may be approximated by using the sample autocovariance given by:

$$c_{xx}(j) \approx \hat{\alpha}_1 c_{xx}(j-1) + \hat{\alpha}_2 c_{xx}(j-2) + \ldots + \hat{\alpha}_m c_{xx}(j-m), \tag{4.15}$$

$j = 1, 2, \ldots, m$. Approximate estimates can be obtained for the parameters $\alpha_1, \alpha_2, \ldots, \alpha_m$, by solving the m simultaneous equations (4.15).

The residual sum of squares and the residual variance may be obtained by using the following approximations:

$$S(\hat{\mu}, \hat{\alpha}_1, \ldots, \hat{\alpha}_m) \approx (n-m) \left\{ c_{xx}(0) - \hat{\alpha}_1 c_{xx}(1) - \ldots - \hat{\alpha}_m c_{xx}(m) \right\} , \tag{4.15a}$$

and

$$S_z^2 \approx \frac{1}{n-2m-1} \; S(\hat{\mu}, \hat{\alpha}_1, \ldots, \hat{\alpha}_m) , \tag{4.15b}$$

respectively.

4.3 Checking The Fit of The Model

Once we have selected the best process that characterizes the ionospheric data and have its parameters estimated, diagnostic checks are made on the model to determine its adequacy. Using this model, we can obtain a series that should simulate the behavior of the original soundings. If the original soundings were filtered, that is, d was different from zero, it would now be necessary to use a "backwards filter," replacing y_t in the model with $(1-B)^d x_t$, and using the resulting process to forecast the oblique or vertical incidence soundings. For example, if we fitted a first-order autoregressive model:

$$\hat{y}_t - \hat{\mu} = \hat{\alpha}_1 (y_{t-1} - \hat{\mu}) + Z_t , \tag{4.16}$$

where $y_t = (1-B)x_t = x_t - x_{t-1}$ is the filter used in the original soundings, then inserting the filter into (4.16), we have:

$$x_t = \varphi_0 + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + Z_t ,$$

where $\varphi_0 = \hat{\mu}(1-\hat{\alpha}_1)$, $\varphi_1 = (1+\hat{\alpha})$ and $\varphi_2 = -\hat{\alpha}_1$. Thus, the mth order autoregressive process, using a first difference filter, can be written as follows:

$$x_t = \varphi_0 + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \ldots + \varphi_{m+d} \, x_{t-m-d} + Z_t , \tag{4.17}$$

where the values of φ_i, $i = 1, 2, \ldots, m+d$, will depend on m and d.

For the fitted model to give a good characterization of the iono-spheric data, the residuals, $r_t = x_t - \hat{x_t}$, $t = 1, 2, \ldots, n$, should behave approximately like random deviates. Hence, the sample autocorrelation function should effectively be zero for all lags except the zero[th] lag.

4.4 Forecasting and Updating the Model

One of the aims in having fitted an autoregressive process to the ionospheric data, is to forecast future values of the oblique or vertical incidence critical frequencies. If we wish to forecast a particular iono-spheric sounding, $x_{t+\ell}$, $\ell \geq 1$, when we are presently at time slot t, then the forecast is made at origin t for a lead-time ℓ. Of course, the shorter the lead-time ℓ, the more accurate our forecasted value will be.

The minimum mean square error forecast for any lead time is given by the conditional expectation, [1], $E_t[X_{t+\ell}]$, of $x_{t+\ell}$, at time slot (origin) t, given knowledge of all x's up to time t. That is,

$$E_t[X_{t+\ell}] = x_t(\ell)$$

Replacing t with $t + \ell$ in equation (4.17), we have:

$$x_{t+\ell} = \varphi_0 + \varphi_1 x_{t+\ell} + \varphi_2 x_{t+\ell-2} + \cdots + \varphi_{n-d} x_{t+\ell-n-d} + Z_{t+\ell} .$$

The minimum mean square error forecast of the ionospheric data is given by:

$$E_t[X_{t+\ell}] = \varphi_0 + \varphi_1 E_t[X_{t+\ell-1}] + \cdots + \varphi_{n+d} E[X_{t+\ell-n-d}] + E[Z_{t+\ell}] . \tag{4.18}$$

For j, a non-negative integer, we know, [1], that:

$$E_t[X_{t+j}] = \hat{x_t}(j), \quad E_t[Z_{t+j}] = 0, \quad j = 1, 2, \ldots , \tag{4.19}$$

and

$$E_t[X_{t-j}] = x_{t-j}, \quad E_t[Z_{t-j}] = Z_{t-j} = x_{t-j} - \hat{x}_{t-j-1} . \tag{4.20}$$

Therefore, we can write equation (4.13) as follows:

$$\hat{x_t}(\ell) = \varphi_0 + \varphi_1 x_{t+\ell-1} + \cdots + \varphi_{n+d} x_{t+\ell-n-d} . \tag{4.21}$$

The variance of the ℓ step ahead forecast-error for any time slot t, is the expected value of:

$$\mathcal{E}_t^2(\ell) = [X_{t+\ell} - \hat{x_t}(1)]^2 . \tag{4.22}$$

Box and Jenkins, [2], have shown that the variance of the lead time, ℓ, is given by:

$$\text{Var}(\ell) = \left\{ 1 + \sum_{j=1}^{\ell-1} \theta_j^2 \right\} \sigma^2 \tag{4.23}$$

where σ_z^2 is estimated by s_z^2, that is:

$$s_z^2 = \frac{S(\hat{\mu}, \hat{\alpha}_0, \ldots, \hat{\alpha}_m)}{n} \; ,$$

and θ_j is given by:

$$\theta_j = 0, \; j < 0$$

$$\theta_0 = 1$$

$$\theta_1 = \varphi_1$$

$$\theta_2 = \varphi_1 \theta_1 + \varphi_2$$

$$\cdot$$
$$\cdot$$
$$\cdot$$

$$\theta_j = \varphi_1 \theta_{j-1} + \cdots + \varphi_{m+d} \, \theta_{j-m-d} \; . \tag{4.24}$$

The $(1-\alpha)\%$ confidence limits for $x_{t+\ell}$ is given by:

$$Pr \left\{ \hat{x}_t(\ell) - U_{\frac{\alpha}{2}} (1 + \sum_{j=1}^{\ell-1} \theta_j^2)^{\frac{1}{2}} s_z \leq x_{t+\ell} \leq \hat{x}_{t+\ell} + U_{\frac{\alpha}{2}} (1 + \sum_{j=1}^{\ell-1} \theta_j^2)^{\frac{1}{2}} s_z \right\} = 1 - \alpha$$

where $U_{\frac{\alpha}{2}}$ is the deviate from the unit normal probability distribution.

In ionospheric problems, we are often interested in forecasting future values of an observed series for several time slots in advance. When we forecast values at lead ℓ greater than or equal t two ($\ell \geq 2$) with an autoregressive process, the forecasted value will be d endent on previously forecasted values; but, as additional ionospheric data becomes available, we can update our old forecast by:

$$\hat{x}_{t+1}(\ell) = x_t(\ell+1) + \theta_\ell Z_{t+1} \; .$$

That is, the "t" origin forecast of $x_{t+\ell+1}$ can be updated to become the "t + 1" origin forecast of the same value, $x_{t+\ell+1}$, by adding a constant multiple of the one-step ahead forecast error Z_{t+1}, where:

$$Z_{t+1} = X_{t+1} - \hat{x}_t(\ell)$$

is used with multiplier θ_ℓ.

5.0 AN AUTOREGRESSIVE MODEL FOR FORECASTING OBLIQUE INCIDENCE IONOSPHERIC SOUNDINGS OVER A 60KM PATH

In this section, we shall utilize the fitting procedure of Section 4.0 to develop an autoregressive model for forecasting the 6th day oblique incidence (OI) critical frequencies for the 60Km experiment. More specifically, we have available 85 observed values of oblique incidence critical frequency data as shown in Figure 5.1. We shall proceed by identifying the data as a non-stationary stochastic realization, fit the model, conduct a diagnostic check of the process, and present both a forecasting and updating scheme.

5.1 Identifying The Observed Data

We plotted the data x_t, $t=1, 2, \ldots, 85$ to attempt to visually detect any trend or non-randomness (see Figure 5.1). The graph of the data appears to exhibit non-stationary properties. As a further aid in identifying the data, we calculated the sample autocorrelation function of x_t and conducted statistical tests for trend. The sample autocorrelation function, $c_{xx}(k)$, was calculated for the original data and for the first and second difference filters. These calculations are shown in Tables 5.2, 5.3, and 5.4, respectively.

It is clear that the sample autocorrelation function of the second-difference data dampens out fairly rapidly, indicating that the filtered series, w_t, have reached statistical equilibrium. Furthermore, we performed statistical tests using Kendall's tau test and found that at the 5% level of significance, the original data contained trend. That is, for the first difference data, y_t, the test statistic, z_θ, was found to be -6.529; and for the second difference filter, it was -0.578, which indicates trend at the 5% level of significance ($z_{.05} = \pm 1.645$).

Therefore, the 6th day OI critical frequencies recorded for the 60Km experiment constitute a non-stationary time series. A second difference filter will transform the data into statistical equilibrium.

5.2 Fitting the Autoregressive Model

Using the filtered series w_t, $t=1, 2, \ldots, 85$, we shall fit an autoregressive process. Recall that in order to fit such a model, it is necessary to estimate the parameters for processes of different orders, and then compute the residual variances for each order. With this information, one can decide on the order of the autoregressive process which best fits the recorded data. Using the maximum likelihood equations approximation, (4.15), we calculated estimates of the parameters, the residual sum of squares, (4.15a), and the residual variance, (4.15b), of the soundings for the autoregressive processes of orders 1, 2, ..., 10. Figure 5.5 shows the residual variance plotted against the order and also shows that the minimum residual variance corresponds to an autoregressive process of order two. Therefore, a second order autoregressive process will give the best fit to the filtered 6th day OI data. Using the procedure discussed in Section 4.0, the estimates of the true state-of-nature of the parameters of such a model were found to be:

$$\widehat{\mu} = \sum_{t=1}^{83} \frac{w_t}{83} = 0.004 ,$$

$$\widehat{\alpha}_1 = 0.409 ,$$

and

$$\widehat{\alpha}_2 = 0.255 .$$

Thus, the autoregressive model for the filtered 6th day data is:

$$w_t - .004 = -.409 (w_{t-1} - .004) - .255(w_{t-2} - .004) + z_t \qquad (5.1)$$

Note that the parameter estimates satisfy the condition that they must lie within a unit circle.

5.3 Diagnostic Check of the Autoregressive Model

Inserting the backwards filter:

$$x_t - 2x_{t-1} + x_{t-2} = w_t$$

into the filtered model (5.1), we have:

$$\widehat{x}_t - 2x_{t-1} + x_{t-2} - .004 = -.409(x_{t-1} - 2x_{t-2} + x_{t-3} - .004)$$

$$-.255(x_{t-2} - 2x_{t-3} + x_{t-4} - .004) + z_t \qquad (5.2)$$

Simplifying equation (5.2) we obtain the forecasting model for the 6th day OI data:

$$\widehat{x}_t = .006 + 1.591 x_{t-1} - .43 x_{t-2} + .1x_{t-3} - .255 x_{t-4} + z_t \qquad (5.3)$$

For simplicity, we write equation (5.3) as follows:

$$\widehat{x}_t = \varphi_0 + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \varphi_3 x_{t-3} + \varphi_4 x_{t-4} + z_t \qquad (5.4)$$

where:

$$\varphi_0 = \widehat{\mu}(1 - \widehat{\alpha}_1 - \widehat{\alpha}_2) = .006$$

$$\varphi_1 = 2 + \widehat{\alpha}_1 = 1.591$$

$$\varphi_2 = \widehat{\alpha}_2 - 2\widehat{\alpha}_1 - 1 = -.436 \qquad \Big\} \quad (5.5)$$

$$\varphi_3 = \widehat{\alpha}_1 - 2\widehat{\alpha}_2 = .1$$

$$\varphi_4 = \widehat{\alpha}_2 = -.255$$

To simulate x_t, we use (5.5) set the unknown value z_t equal to its unconditional expectation of zero, and assume the values x_{t-1}, x_{t-2}, x_{t-3}, and x_{t-4} are known. Figure 5.6 shows the simulation of the observed series,

x_t, t=1, 2, ..., 85. We obtain the residuals by subtracting the modeled series from the observed series, that is:

$$r_t = x_t - \hat{x}_t, \quad t=5, \ldots, 85 \tag{5.6}$$

The residuals start at time slot 5, and $r_1 = r_2 = r_3 = r_4 = 0$. That is, it is necessary that the first four values of the data be known and that forecasting begins at the next time slot. The residual sum of squares was found to be:

$$\sum_{t=5}^{85} r_t^2 = 1.994 \tag{5.7}$$

The behavior of the residuals approximates that of a purely random process, where their sample autocorrelation function should be effectively zero. Figure 5.7 shows that we have developed a good model for the ionospheric soundings over a 60Km path.

Since n is sufficiently large, the sample autocorrelation function, $r_{zz}(k)$ is approximately normal with mean zero and variance $1/n$, [3]. The standard deviation of $r_{zz}(k)$ was found to be 0.11 and the 95% confidence limits for $r_{zz}(k)$ of the residuals is:

$$r_{zz}(k) \pm 0.11 \ (1.96) \ ,$$

or:

$$Pr[r_{zz}(k)-.22 \leq \rho_{zz}(k) \leq r_{zz}(k) + .22] = .95,$$

or see Figure 5.7.

5.4 Forecasting and Updating

The fitted model:

$$\hat{x}_t = \varphi_0 + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \varphi_3 x_{t-3} + \varphi_4 x_{t-4} + z_t$$

may be used to forecast future values of the observed series $\hat{x}_t$. To forecast ahead for a lead time, ℓ, we have:

$$\hat{x}_{t+\ell} = \varphi_0 + \varphi_1 x_{t+\ell-1} + \varphi_2 x_{t+\ell-2} + \varphi_3 x_{t+\ell-3} + \varphi_4 x_{t+\ell-4} + z_{t+\ell} \ .$$

The minimum mean square error forecast is given by:

$$\hat{x}_t(\ell) = \varphi_0 + \varphi_1 x_{t+\ell-1} + \varphi_2 x_{t+\ell-2} + \varphi_3 x_{t+\ell-3} + \varphi_4 x_{t+\ell-4} \ .$$

To illustrate how one may update the forecasts for a slot time t, (origin) suppose that a new piece of data, x_{t+1}, becomes available. With the origin at time slot t+1, we update the forecast by:

$$\hat{x}_{t+1}(\ell) = \hat{x}_t(\ell+1) + \theta_\ell z_{t+1}, \quad \ell = 1, 2, \ldots, 10$$

where: $z_{t+1} = x_{t+1} - \hat{x}_t(1)$ and θ_ℓ is as described in Section 4.0.

Table 5.8 shows the forecasted values of the 6th day oblique incidence series for 1, 2, ..., 11 time slots ahead at origin t=59 along with the 50% and 95% probability limits. The actual values of x_t are shown for comparison, but these values are not actually known when the forecast is made. It also shows the updated forecast for 1, 2, ..., 10 time slots ahead at origin t=60. Note that the updated forecast is an improvement over the original forecast at every lead time.
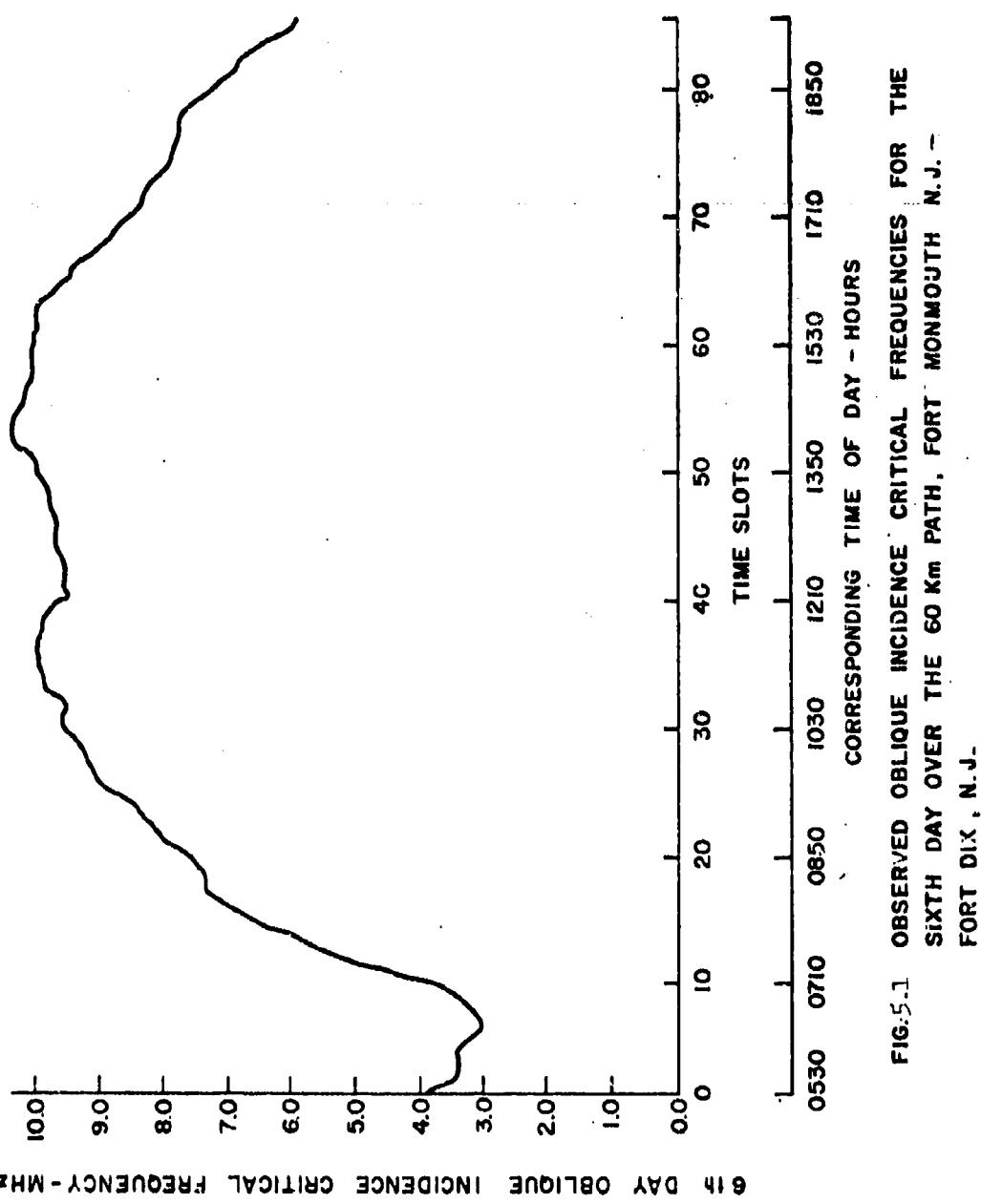
FIG. 5.1 OBSERVED OBLIQUE INCIDENCE CRITICAL FREQUENCIES FOR THE SIXTH DAY OVER THE 60 Km PATH, FORT MONMOUTH N.J. — FORT DIX, N.J.

Sample Auto Correlation Of The 6th Day Oblique Incidence
Critical Frequencies For The 60 Km Path, Fort Monmouth, N.J. -
Fort Dix, N.J.

Lags										
1-10	.96	.92	.86	.81	.74	.68	.61	.54	.47	.41
11-20	.35	.52	.25	.20	.16	.12	.08	.04	.00	-.04
21-30	-.07	-.10	-.12	-.15	-.17	-.18	-.20	-.21	-.22	-.23
31-40	-.24	-.25	-.26	-.26	-.26	-.27	-.27	-.28	-.29	-.30
41-50	-.31	-.33	-.34	-.35	-.35	-.36	-.36	-.36	-.36	-.35
51-60	-.34	-.32	-.31	-.29	-.27	-.25	-.23	-.20	-.18	-.15
61-70	-.12	-.09	-.06	-.04	-.01	.01	.03	.05	.07	.09
71-80	.10	.11	.12	.13	.14	.14	.14	.13	.12	.11
81-84	.08	.07	.05	.02						

TABLE 5.2

Sample Auto Correlation Of The First Difference Data
Of The 6th Day Oblique Incidence Critical Frequencies
For The 60 Km Path, Fort Monmouth, N.J. – Fort Dix, N.J.

Lags										
1-10	.67	.58	.56	.45	.33	.25	.18	.18	.21	.19
11-20	.23	.17	.21	.20	.14	.16	.11	.11	.06	.03
21-30	.01	.03	-.02	.01	-.01	-.02	-.03	-.04	-.04	-.04
31-40	-.06	-.06	.02	-.02	.01	.00	.02	.02	-.02	-.02
41-50	-.03	-.06	-.10	-.12	-.13	-.17	-.18	-.16	-.15	-.22
51-60	-.19	-.18	-.22	-.25	-.24	-.27	-.26	-.24	-.23	-.22
61-70	-.15	-.14	-.13	-.14	-.09	-.10	-.14	-.16	-.19	-.18
71-80	-.15	-.16	-.14	-.06	-.03	.00	.05	.07	.06	.03
81-83	.05	.04	.92							

TABLE 5.3

Sample Auto Correlation Of The Second Difference Data Of The
6th Day Oblique Incidence Critical Frequencies For The 60 Km
Path, Fort Monmouth, N.J. - Fort Dix, N.J.

| Lags										
1-10	-.33	-.12	.12	.02	-.01	-.01	-.08	.01	.10	-.11
11-20	.09	-.12	.07	.05	-.11	.05	-.05	.08	.02	-.07
21-30	-.06	.13	-.09	.04	-.02	.00	.32	-.04	-.03	.07
31-40	.00	-.17	.22	-.13	.03	-.01	-.01	.05	-.02	.01
41-50	.05	.02	-.06	-.01	.04	-.04	-.05	.02	.13	-.10
51-60	-.04	.09	-.05	-.04	.06	-.04	-.02	.08	.00	-.06
61-70	.06	-.02	.03	-.06	.04	.07	-.04	.05	-.07	-.02
71-80	-.06	-.06	-.04	.05	.00	-.05	.01	.04	.02	-.07
81-82	.00	.04								

TABLE 5.4
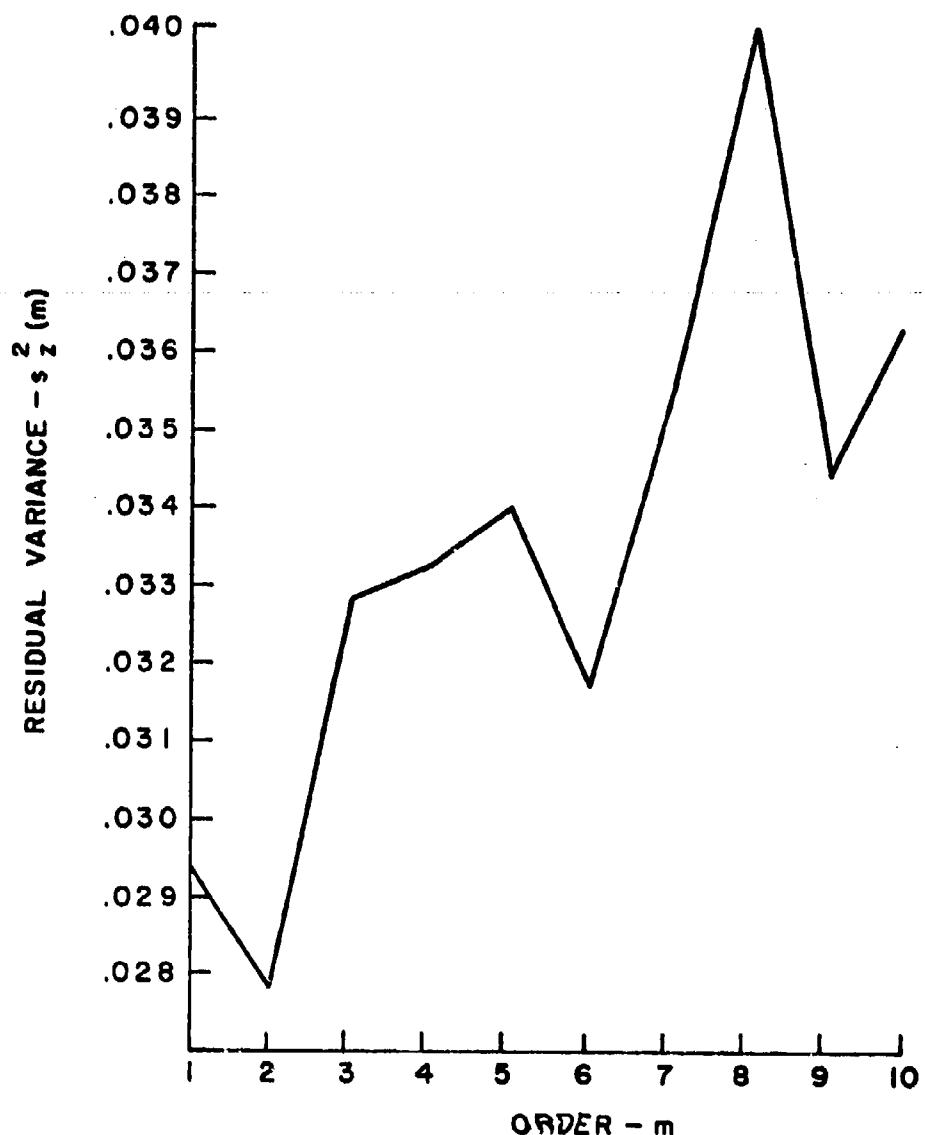
FIG. 5.5 RESIDUAL VARIANCE OF THE AUTOREGRESSIVE MODEL, ORDERS 1-10, OF THE 6 th DAY OBLIQUE INCIDENCE CRITICAL FREQUENCIES FOR THE 60 Km PATH, FORT MONMOUTH, N.J. – FORT DIX, N.J.

FILTERED MODEL:

$$\hat{W}_t - .004 = Z_t - .409 (W_{t-1} - .004) - .255 (W_{t-2} - .004)$$

AUTOREGRESSIVE FORECASTING MODEL:

$$\hat{X}_t = .006 + 1.591 X_{t-1} - .436 X_{t-2} + .10 X_{t-3} - .255 X_{t-4}$$

—— OBSERVED

----- FORECASTED

TIME SLOTS

CORRESPONDING TIME OF DAY - HOURS

6 TH DAY OBLIQUE INCIDENCE CRITICAL FREQUENCY - MHz
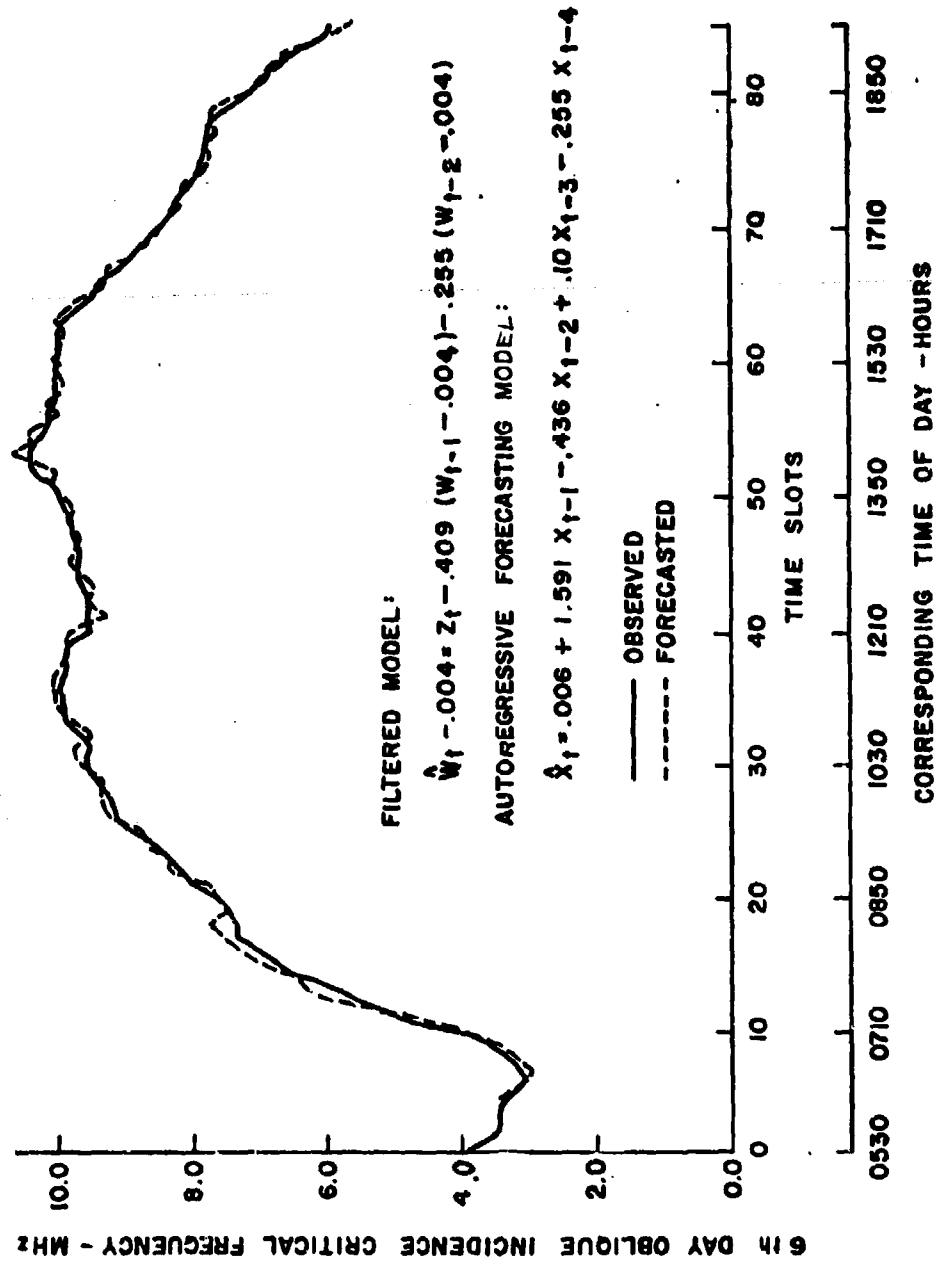
FIG.5.6 FORECASTED SERIES, USING THE AUTOREGRESSIVE MODEL VS OBSERVED SERIES OF THE OBLIQUE INCIDENCE CRITICAL FREQUENCIES FOR THE SIXTH DAY OVER THE 60 Km PATH, FORT MONMOUTH, N.J. - FORT DIX, N.J.
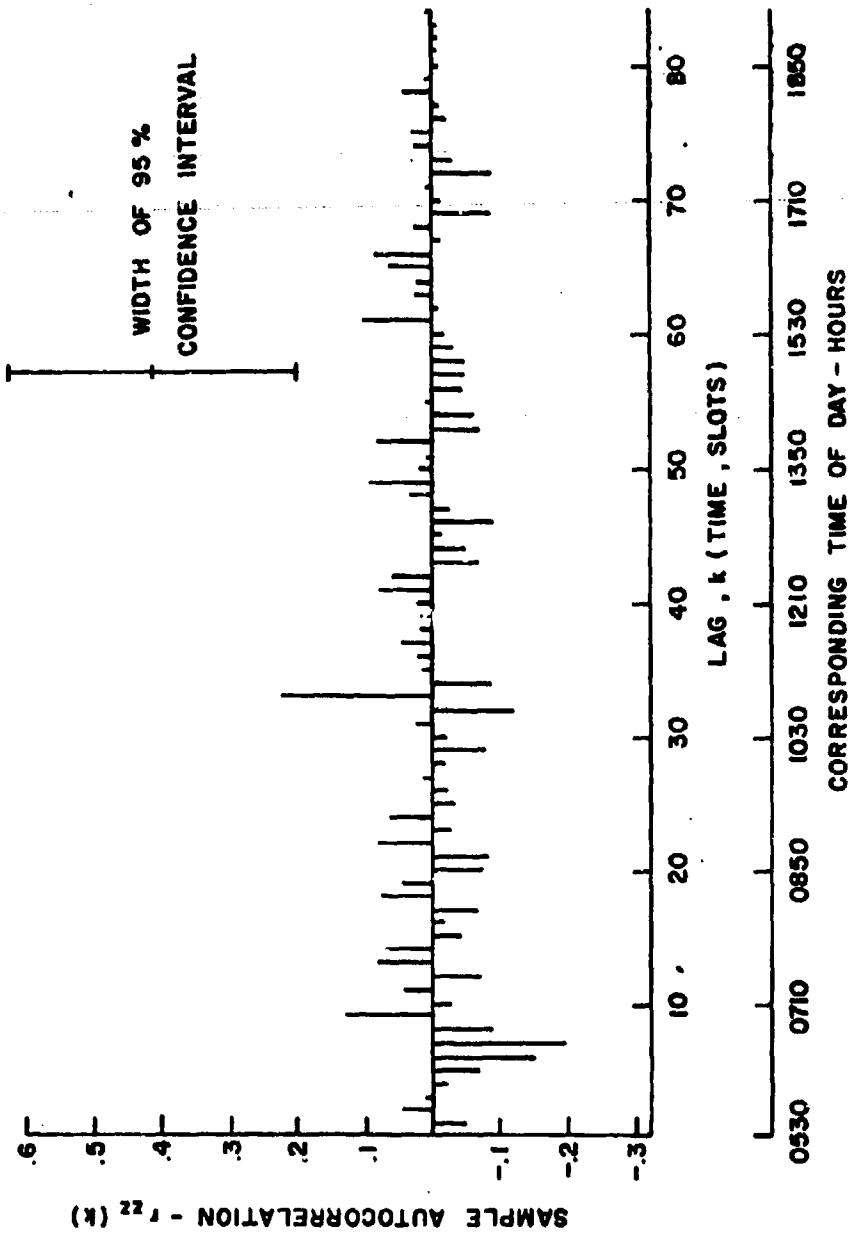
FIG.5.7 SAMPLE AUTOCORRELATION FUNCTIONS AND 95 % CONFIDENCE
INTERVAL OF THE RESIDUALS OF THE OBLIQUE INCIDENCE CRITICAL
FREQUENCIES FOR THE SIXTH DAY OVER THE 60 Km PATH,
FORT MONMOUTH, N.J. -FORT DIX, N.J

Forecasted Values With 50% and 95% Confidence Intervals Of The
6th Day Observed Oblique Incidence Critical Frequencies For The
60 Km Path, Fort Monmouth, N.J. - Fort Dix, N.J., For Lead Times
Of 1,2,...,11 Steps Ahead At Origin T = 59, And The Updating Of
These Forecasts Assuming $\bar{x}_{60}$ Becomes Available Using The Fitted
Autoregressive Model

TIME	ACTUAL VALUE	LEAD TIME	FORECASTED VALUE	PROBABILITY LIMITS 50%	95%	LEAD TIME	UPDATED FORECAST
60	10.001	1	10.043	±0.189	±0.317		
61	9.961	2	10.042	±0.206	±0.598	1	9.991
62	10.000	3	10.054	±0.307	±0.892	2	9.987
63	9.911	4	10.065	±0.429	±1.249	3	9.977
64	9.650	5	10.079	±0.563	±1.639	4	9.972
65	9.450	6	10.098	±0.708	±2.058	5	9.973
66	9.411	7	10.121	±0.863	±2.511	6	9.976
67	9.100	8	10.147	±1.031	±2.997	7	9.983
68	8.833	9	10.176	±1.206	±3.508	8	9.993
69	8.678	10	10.210	±1.391	±4.054	9	10.008
70	8.489	11	10.247	±1.583	±4.504	10	10.026

TABLE 5.8

6. A MOVING AVERAGE MODEL FOR FORECASTING MEAN VERTICAL INCIDENCE SOUNDINGS FOR THE 60 KM EXPERIMENT

In this section, we shall illustrate the fitting procedure discussed in Section 4 by formulating a difference equation for the mean vertical incidence soundings for the 60 Km experiment. More specifically, we shall fit a moving average model to the 85 observed values of the vertical incidence data, each of which is the mean of nine observed soundings taken at specific time slots during the period of the experiment.

6.1 Identifying The Series:

A plot of the data, x_t, t=1, 2, ..., 85, is shown in Figure 6.1. The visual interpretation of the time series is that it exhibits non-stationary properties. To substantiate this, we calculated the sample auto-correlation function. Table 6.2 shows that the sample autocorrelation does not dampen out very rapidly. Furthermore, Kendall's tau test, [5], for trend, was applied and, at the 5% level of significance, we confirmed the fact that the ionospheric data was not in statistical equilibrium.

A first difference filter:

$$y_t = x_t - x_{t-1}, \ t = 1, 2, ..., 85,$$

and a second difference filter:

$$w_t = x_t - 2x_{t-1} + x_{t-2}, \ t = 1, 2, ..., 85,$$

were applied to the original data. Using the second order difference filter, the sample autocorrelation function, shown in Table 6.4, dampens out more rapidly than if the first difference filter (shown by Table 6.3) were used. Also, the values of Kendall's tau are -6.274 and -0.539 for the first and second difference filters, respectively. Thus, at the 5% level of significance ($z_{.05} = \pm 1.645$), the second difference data indicates that the VI soundings are in statistical equilibrium.

6.2 Fitting the Moving Average Model:

Using the maximum likelihood method, we estimated the parameters of the moving average model for orders up to five, and then we computed the residual variance for each order. As indicated above, the criterion for selecting the order of the process which gives the best fit to the mean vertical incidence data will be the order which has minimum residual variance.

Figure 6.5 shows that the minimum residual variance for the moving average process will be given by order two. Hence, a second-order moving average process will give the best fit to the filtered series, w_t. The parameters estimated for this model were found to be:

$$\hat{\mu} = 0.0 \ , \ \hat{\beta}_1 = 0.58 \ , \ \text{and} \ \hat{\beta}_2 = -0.12$$

Note that the invertibility conditions discussed above are satisfied by the estimates of the parameters. Thus, the moving average process that characterizes the filtered data is:

$$\hat{w}_t = z_t - .58\ z_{t-1} + .12\ z_{t-2}\ .$$

6.3 Diagnostic Check Of The Model:

To determine the adequacy of the fitted process, we must simulate the observed series, and then calculate the residuals to see if they behave as a purely random process. Recall that, in order to use the above model, $\hat{w}_t$, to simulate the observed series, x_t, we must make use of the "backwards filter" which depends on the original filter employed to transform the observed soundings. Hence, inserting the backward filter:

$$x_t - 2x_{t-1} + x_{t-2} = w_t$$

to our model, $\hat{w}_t$, we have:

$$\hat{x}_t = 2x_{t-1} - x_{t-2} + z_t - .58\ z_{t-1} + .12\ z_{t-2}\ .$$

Setting the unknown values of z_t equal to their unconditional expectation of zero, we begin the simulation by initially assuming x_1 and x_2 are known; to simulate x_t, we assume x_{t-1} and x_{t-2} are known. Figure 6.6 shows the simulated series along with the original data.

To check if the residuals behave as a purely random process, we calculated the sample autocorrelation function, $r_{zz}(k)$, of the residuals, for lags $k = 1, 2, \ldots, 84$. Figure 6.7 shows that $r_{zz}(k)$ indeed dampens out rapidly.

Since the total number of ionospheric soundings is sufficiently large, $n = 85$, $r_{zz}(k)$ is approximately Gaussian with mean zero and variance $1/n$. The calculated value of the standard deviation of $r_{zz}(k)$ was found to be 0.11 and the 95% confidence limits are:

$$r_{zz}(k) \pm .22\ .$$

Figure 6.7 shows that none of the sample autocorrelations of the residuals are outside the above confidence limits.

6.4 Forecasting And Updating:

To forecast ahead "ℓ" slots with the above moving average model, we have:

$$\hat{x}_{t+\ell} = 2x_{t+\ell-1} - x_{t+\ell-2} + z_{t+\ell} - .58z_{t+\ell-1} + .12z_{t+\ell-2}\ .$$

The minimum mean square error forecast is given by:

$$\hat{x}_t(\ell) = 2x_{t+\ell-1} - x_{t+\ell-2} - .58z_{t+\ell-1} + .12z_{t+\ell-2}\ .$$

Using $\hat{x}_t(\ell)$ to forecast lead times greater than two, the model becomes:

$$\hat{x}_t(\ell) = 2x_{t+\ell-1} - x_{t+\ell-2} \, , \quad \ell > 2 \, .$$

Note that the forecasted values will not depend on previous errors, z_t.

To update the forecast made at origin t, we assume that a new ionospheric sounding, x_{t+1}, has been realized. Now, with the origin at t+1, we may update the forecast using:

$$\hat{x}_{t+1}(\ell) = \hat{x}_t(\ell+1) + \theta_\ell [x_{t+1} - \hat{x}_t(1)] \, ,$$

where the θ_ℓ are as given in Section 4.

Table 6.8 shows the forecasted values of the data for 1, 2, ..., 5 time slots ahead at origin, t = 40, along with the associated 50% and 95% confidence limits. The actual values of the observed series are shown for comparison, but they were not actually known when the forecast was made. The table also shows the updated forecasts for 1, 2, 3, and 4 time slots in the future at origin t = 41.
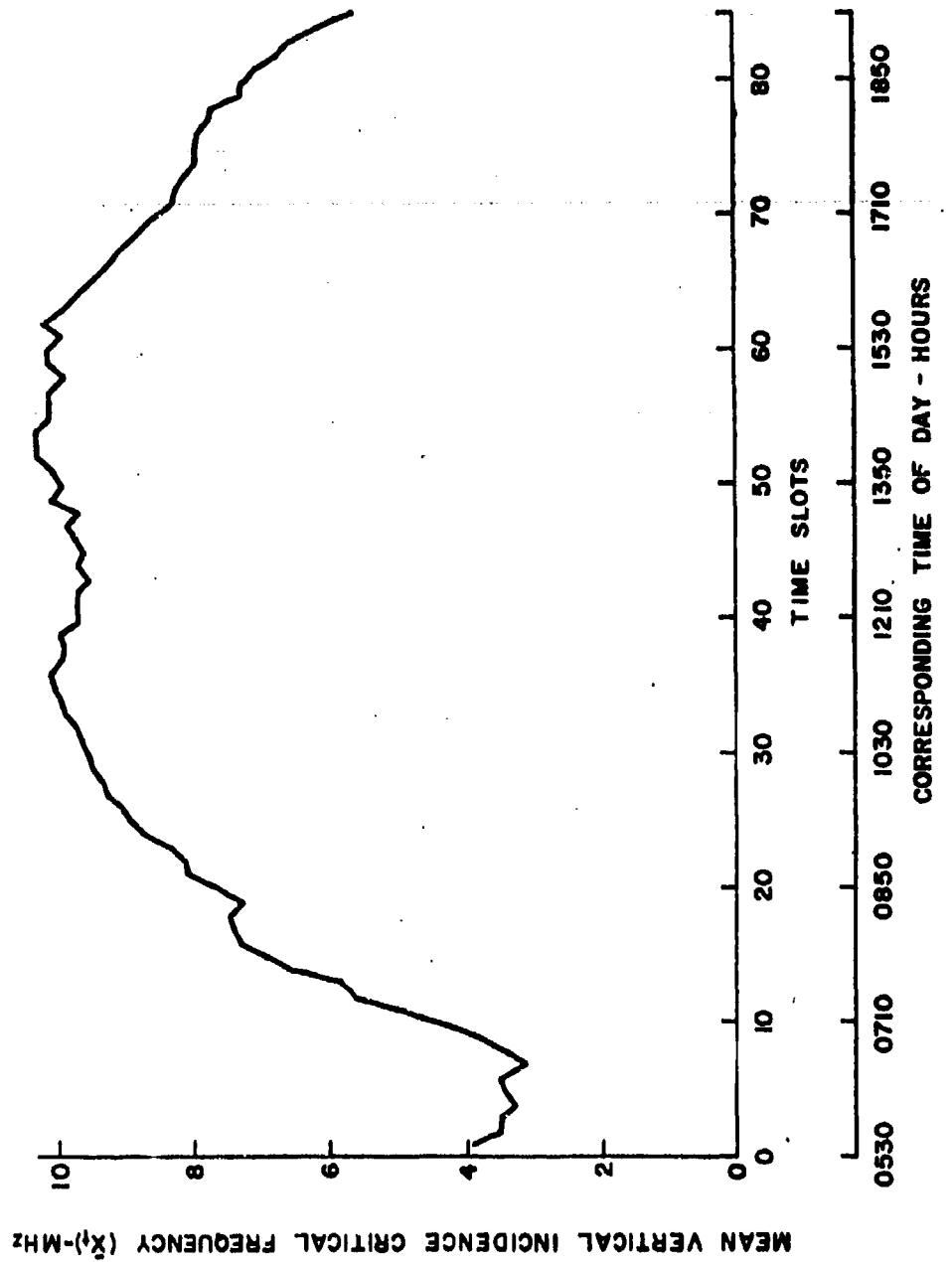
FIG. 6.7 OVERALL AVERAGE OF THE OBSERVED VERTICAL INCIDENCE CRITICAL FREQUENCIES, FOR 9 DAYS, TAKEN AT FORT MONMOUTH, N.J. FOR THE 60 Km EXPERIMENT.

Sample Auto Correlation Of The Overall Means (9Days) Of
The Vertical Incidence Critical Frequencies Taken At
Fort Monmouth, N.J. For The 60 Km Experiment.

Lags										
1-10	.96	.91	.86	.80	.74	.67	.60	.53	.47	.41
11-20	.35	.30	.25	.20	.16	.12	.08	.04	.00	-.03
21-30	-.06	-.10	-.12	-.14	-.16	-.18	-.20	-.21	-.23	-.24
31-40	-.25	-.25	-.26	-.26	-.26	-.27	-.27	-.28	-.29	-.30
41-50	-.31	-.32	-.33	-.34	-.35	-.36	-.36	-.36	-.36	-.35
51-60	-.34	-.32	-.31	-.29	-.27	-.25	-.23	-.21	-.18	-.15
61-70	-.12	-.10	-.06	-.04	-.01	.01	.03	.05	.07	.08
71-80	.10	.11	.12	.13	.11	.11	.11	.13	.12	.11
81-84	.10	.08	.06	.03						

TABLE 6.2

Sample Auto Correlation Of The First Difference Data
Of The Overall Means Of Vertical Incidence Critical
Frequencies Taken At Fort Monmouth, N.J. For The
60 Km Experiment.

Lags										
1-10	.57	.53	.51	.35	.39	.29	.16	.16	.11	.24
11-20	.20	.16	.20	.14	.24	.14	.12	.14	.06	.09
21-30	.06	-.02	-.01	.05	-.05	-.03	-.01	-.02	-.04	-.03
31-40	-.03	-.06	-.03	-.03	.02	.01	.01	.01	-.01	-.06
41-50	.00	-.06	-.10	-.11	-.10	-.09	-.13	-.07	-.21	-.19
51-60	-.13	-.20	-.21	-.22	-.26	-.22	-.24	-.24	-.23	-.22
61-70	-.14	-.16	-.11	-.13	-.12	-.07	-.13	-.17	-.17	-.20
71-80	-.18	-.13	-.18	-.13	-.06	-.02	.03	.04	.02	.03
81-83	.05	.05	.05							

TABLE 6.3

Sample Auto Correlation Of The Second Difference
Data Of The Overall Means Of Vertical Incidence
Critical Frequencies Taken At Fort Monmouth, N.J.
For The 60 Km Experiment.

| Lags										
1-10	-.46	.00	.20	-.24	.13	.10	-.12	.03	-.13	.18
11-20	-.09	-.01	.10	-.25	.23	-.09	-.08	.18	-.14	.02
21-30	.10	-.12	.00	.14	-.14	.01	.00	.02	-.03	.00
31-40	.05	-.08	.08	-.09	.02	.06	-.05	.04	.05	-.15
41-50	.15	-.02	-.04	-.02	-.01	.06	-.10	.20	-.15	-.05
51-60	.13	-.06	-.03	.07	-.12	.08	.00	.00	-.05	.03
61-70	.03	-.06	.03	.03	-.07	.09	-.02	-.04	.05	-.01
71-80	-.03	.09	-.08	-.02	.01	-.03	.00	.00	-.01	-.01
81-82	-.01	-.01								

TABLE 6.4

-493-
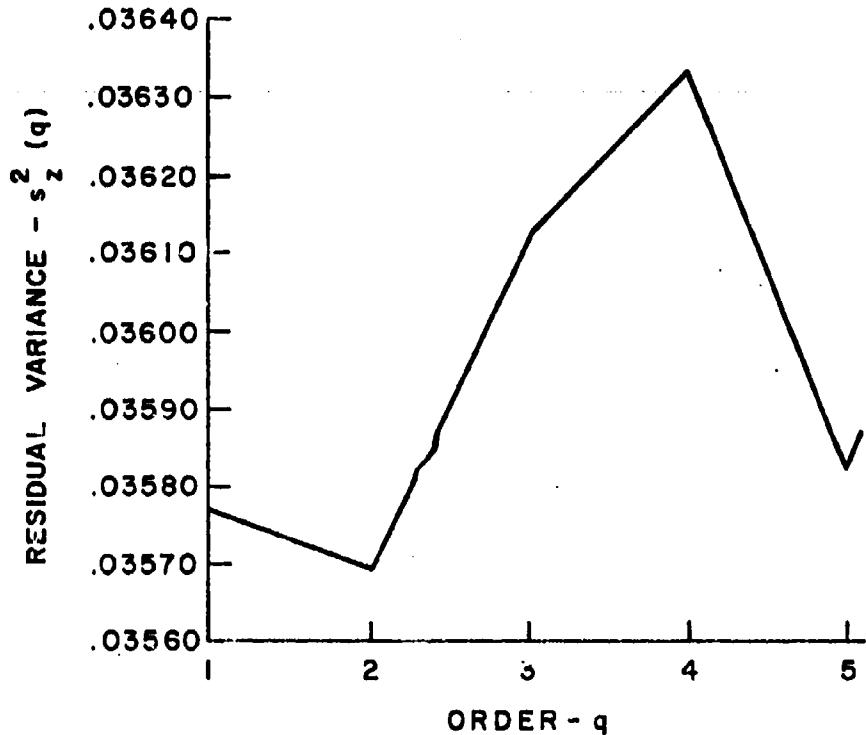
FIG.6.5 RESIDUAL VARIANCE OF THE MOVING AVERAGE MODEL, ORDERS 1-5, OF THE OVERALL MEAN OF VERTICAL INCIDENCE CRITICAL FREQUENCIES TAKEN AT FORT MONMOUTH, N.J. FOR THE 60 Km EXPERIMENT.

FILTERED MODEL:

$$\hat{W}_t = Z_t - .58 Z_{t-1} + .12 Z_{t-2}$$

MOVING AVERAGE FORECASTING MODEL:

$$\hat{X}_t = 2\bar{X}_{t-1} - \bar{X}_{t-2} - .58 Z_{t-1} + .12 Z_{t-2}$$

——— OBSERVED
----- FORECASTED
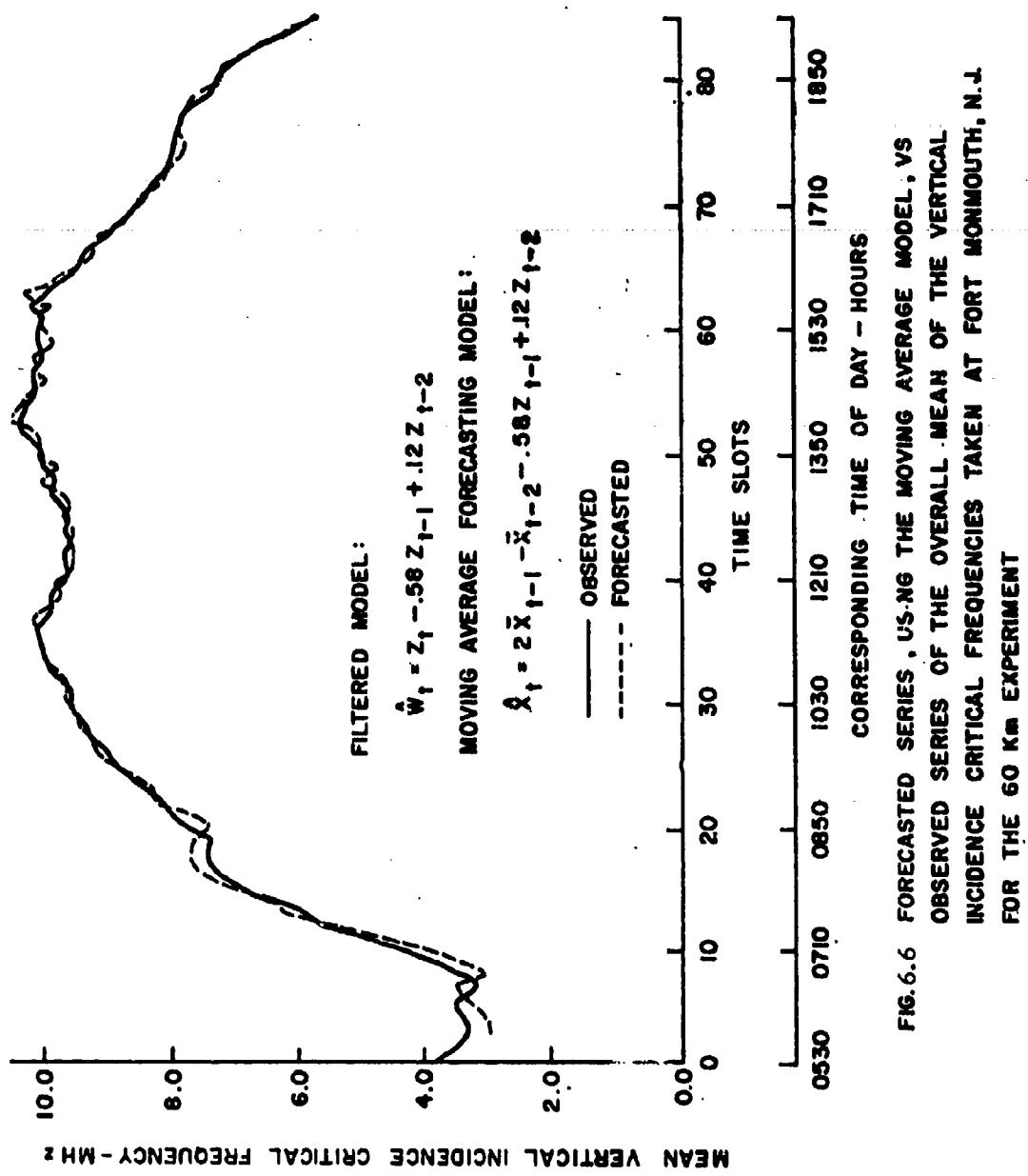
FIG. 6.6 FORECASTED SERIES, USING THE MOVING AVERAGE MODEL, VS OBSERVED SERIES OF THE OVERALL MEAN OF THE VERTICAL INCIDENCE CRITICAL FREQUENCIES TAKEN AT FORT MONMOUTH, N.J. FOR THE 60 Km EXPERIMENT
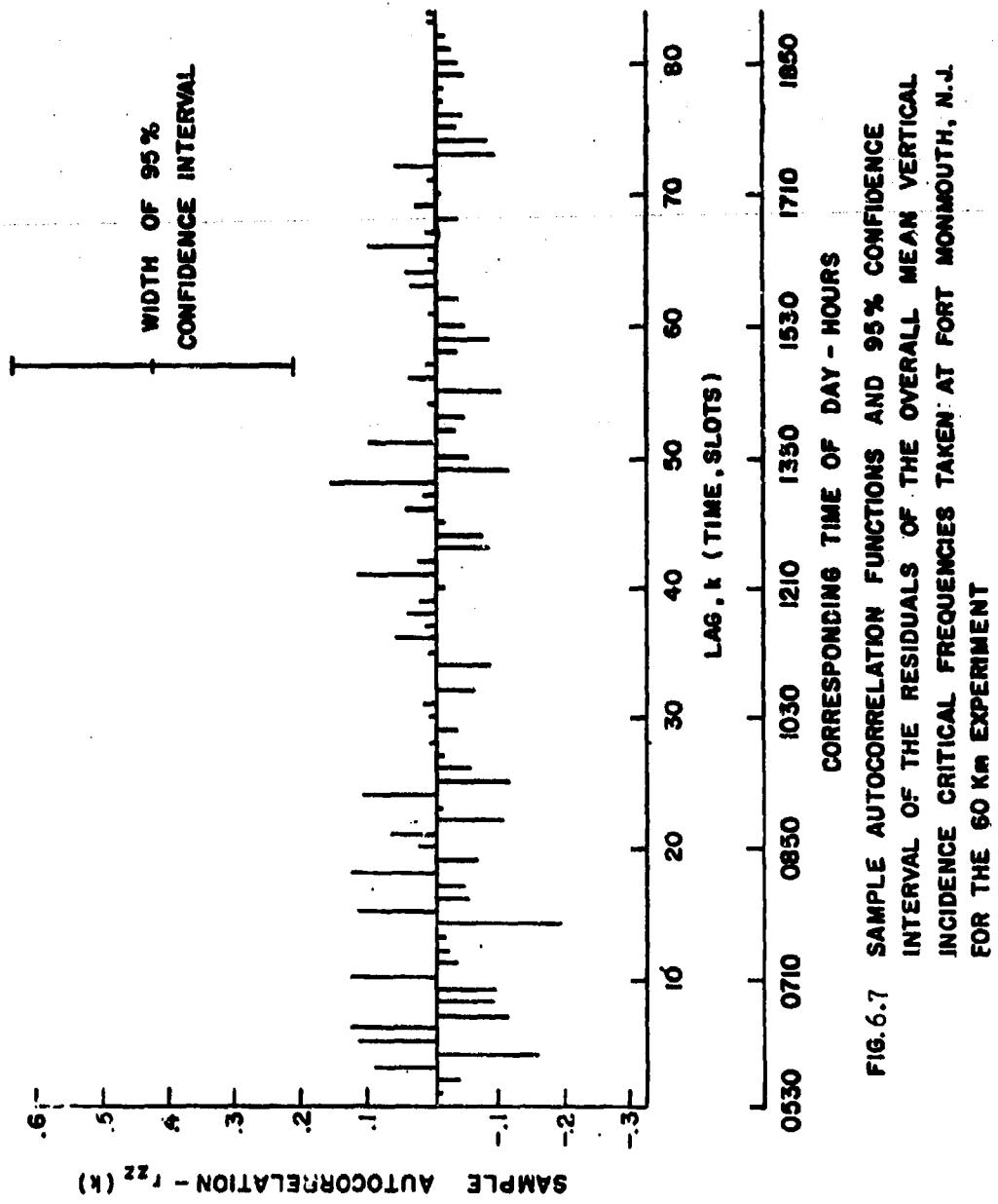
FIG. 6.7 SAMPLE AUTOCORRELATION FUNCTIONS AND 95% CONFIDENCE
INTERVAL OF THE RESIDUALS OF THE OVERALL MEAN VERTICAL
INCIDENCE CRITICAL FREQUENCIES TAKEN AT FORT MONMOUTH, N.J.
FOR THE 60 Km EXPERIMENT

Forecasted Values With 50% and 95% Confidence Intervals Of The
Overall Means Of The Observed Vertical Incidence Critical
Frequencies Taken At Fort Monmouth, N.J. For The 60 Km Experiment
For Lead Times Of 1, 2, ..., 6 Steps Ahead At Origin T = 39, And
The Updating Of These Forecasts Assuming x_{40} Becomes Available
Using The Fitted Moving Average Model.

TIME	ACTUAL VALUE	LEAD TIME	FORECASTED VALUE	PROBABILITY LIMITS 50%	PROBABILITY LIMITS 95%	LEAD TIME	UPDATED FORECAST
40	9.667	1	9.834	± 0.133	±0.357		
41	9.622	2	9.806	± 0.154	±0.414	1	9.709
42	9.656	3	9.778	± 0.157	±0.421	2	9.741
43	9.567	4	9.750	± 0.157	±0.422	3	9.740
44	9.689	5	9.722	± 0.157	±0.422	4	9.721
45	9.600	6	9.694	± 0.157	±0.422	5	9.694

TABLE 6.8

7. SUMMARY AND CONCLUSIONS

Oblique incidence (OI) and vertical incidence (VI) ionospheric soundings were obtained as described in section 2 over 60 Km, 200 Km, and 500 Km distances. The main objectives of these experiments were to develop a forecasting model to predict short-term future ionospheric conditions for these specific distances, knowing VI data at a particular time, and to estimate equivalent OI data over the area of interest. In view of these experimental investigations, over the paths specified, the following results were obtained along with the related conclusions:

(a) It was pointed out that the widely accepted approach (secant law) of estimating equivalent OI from VI data for what are considered tactical Field Army distances (especially at 500 Km), does not give the desired accuracy for practical use within the Field Army area of responsibility.

(b) A table has been formulated (p. 8) which shows that at the three distances investigated, there exists a very strong linear dependence between observed OI and VI data.

(c) In view of the strong linear dependence demonstrated, linear regression models were developed for all three distances. More specifically, two regression models were developed for each of the distances investigated (and also for each of the reciprocal paths), namely, one for a specific day, chosen at random, and one for the overall mean of the observed data.

(d) It has been shown in Section 3, that one can obtain an excellent estimate of the equivalent OI data at a particular time slot, for a given path, knowing the VI information.

(e) For each of the linear regression models developed, we gave 95% confidence intervals. That is, we are 95% certain that these bounds contain the true state of nature which we have estimated.

(f) One could use the graphical presentation of the regression models (pp. 10-21) to estimate equivalent OI data. For the experiments performed, these results can be directly compared to the actual observed OI data.

(g) One of the needs of the tactical Field Army for more effective H. F. communications, is to develop a system to forecast the VI soundings at specific times in advance for specific distances. Having such a forecasted value at a particular time, one can insert it into a regression model to obtain an estimate of the equivalent OI soundings at the same time in advance.

(h) In view of the above need, a precise procedural approach (Sections 5 and 6), utilizing time series analysis techniques, has been given for the development of such forecasting models.

(i) A complete analysis using the above procedure was carried out for the 60 Km experiment. It was shown that both the OI and VI soundings were non-stationary stochastic realizations and filtering of the data was necessary.

-498-

(j) An autoregressive model was developed for forecasting OI soundings over a 60 Km path. This resulted in the following difference equation:

$$\hat{w}_t = .006 + 1.591\, x_{t-1} - .436\, x_{t-2} + .10\, x_{t-3} - .255\, x_{t-4}\ .$$

One can utilize this model to forecast the OI recording at 1, 2, ..., k time slots in advance. It is necessary that we have four (4) initial values of the experiment.

(k) A method has also been given where the model can be updated. That is, if additional information becomes available, it can be utilized to improve the forecasted values for many time slots in advance.

(l) A table has been formulated (p. 41) where we utilized the autoregressive model to forecast OI data up to ten time slots in advance. Also given are 50% and 95% confidence bounds of the true state of nature associated with this physical phenomenon.

(m) A moving average model has been developed for forecasting vertical incidence (VI) ionospheric soundings for the 60 Km experiment at the Fort Monmouth terminal.

(n) As was the case with the OI data, the VI data exhibited non-stationary properties. A second order filter was necessary to transform the data into statistical equilibrium.

(o) The moving average model which characterizes the VI soundings is given by:
$$\hat{\bar{x}}_t = 2x_{t-1} - x_{t-2} - .58z_{t-1} + .12z_{t-2}\ .$$

One can utilize the above model to obtain a VI value at a given time slot in the future, and by using this value in the corresponding linear regression model, we can obtain the necessary future estimate of the equivalent OI sounding.

(p) The sample autocorrelation function of the residuals, that is, the actual value minus the forecasted value, was calculated. It dampened out fairly rapidly, which indicates the effectiveness, in terms of accuracy, of the proposed model.

(q) A table has been given (p. 52) where the moving average model has been utilized to forecast VI data, 1, 2, ..., 6 time slots in advance. Also given are the 50% and 95% confidence intervals.

The findings and the models that have been developed in this study can be utilized by the tactical Field Army for the approximate distance involved under similar circumstances of geographical location, seasonal effects, and geomagnetic activity. This is not to say that the models could not be revised so that they could be universally used for this tactical distance.

The authors are presently involved in formulating statistical forecasting models for the 200 Km and 500 Km experiments.

REFERENCES

1. Box, G. E. P., and Jenkins, G. M., "Some Recent Advances in Forecasting and Control," Applied Statistics, 17, 91 (1968)

2. Box, G. E. P., and Jenkins, G. M., Time Series Analysis Forecasting and Control, Holden-Day, San Francisco, Calif. (1970)

3. Jenkins, G. M., and Watts, D. G., Spectral Analysis and Its Application, Holden-Day, San Francisco, Calif. (1968)

4. Jury, E. I., Theory and Applications of the Z-Transform Method, John Wiley, New York (1964)

5. Kendall, M. G., and Stuart, A., The Advanced Theory of Statistics, Vol.3, Griffin, London, England (1966)

6. D'Accardi, R. J., and Tsokos, C. P., "Regression Modeling of Oblique and Incidence Ionospheric Soundings for Different Paths," U. S. Army Electronics Command Technical Report, Fort Monmouth, N. J. (1972)

7. Krause, G. E., et al, "Field Test of a Near Real-Time Ionospheric Forecasting Scheme, (500 km)," Research and Development Technical Report, U. S. Army Electronics Command, to be Published

SOME PROBLEMS IN THE DESIGN OF
TESTS TO CHARACTERIZE IR BACKGROUND TRANSIENTS

J. S. Dehne; J. R. Schwartz; A. J. Carillo
Combat Surveillance and Target Acquisition Laboratory,
US Army Electronics Command, Fort Monmouth, New Jersey

ABSTRACT

Ultimate development of military weapons flash detection/location equipment is hindered by a lack of engineering data characterizing the AC component of the infrared background.

Experimental determination of the needed data has proved very difficult due to problems in conceiving of an experiment which can separately measure all the AC background components (e.g. background changes are indistinguishable from atmospheric turbulence, etc). Without such separation of effects, it is impossible to reduce results to terms of engineering interest.

Requirements on the needed data and experimental equipment, and procedures are discussed.

INTRODUCTION

Military requirements exist for systems to detect and locate hostile ordinance of various types. In recent years several possible solutions to the problem have been investigated. The detection and location of weapons by their muzzle flash has proved to be the most feasible approach. Recent studies have shown that staring systems, operating in the near infrared, are superior to other proposed approaches to flash detection/location.

The detection capabilities of such systems are limited under daylight conditions by the AC component or transient nature of the background radiation. Several difficulties have been encountered in our initial attempts to characterize this phenomenon.

These difficulties seem to involve two distinct problem areas for which outside aid and guidance is sought. First, we lack a theoretical explanation of the phenomenon under investigation. Without such a theory it is difficult to design definitive measurement experiments. Second, we need advice on available data reduction procedures to insure the generality and utility of our data.

Description of Staring Flash Detection System

A simple flash detection system consists of optics, a photodetector, and associated electronics.

The optics collect light and focus it on the photodetector. The ratio of collecting area to detector area determines the optical gain. The physical dimensions of the detector and the focal length of the optics fixes the field of view of the system. Targets outside this field of view are not imaged on the detector; hence are not detected. An optical filter is used to pass light within a specified wavelength region (the optical passband) and to attenuate all other wavelengths.

The photodetector provides an electrical output which is proportional to the flux density of light (irradiance) impinging on it.

Threshold detection electronics for such systems may be quite complex. The simplest version would include an electronic filter circuit and a threshold detector. The electronic filter is used to optimise the signal to noise ratio. Spurious noise signals generated by the background or the photodetector are to be attenuated as much as possible. Many well known criteria are available to guide filter design if the statistical natures of the signal and noise are known.

The threshold detector passes only those signals which satisfy a given detection criteria. The simplest of these conditions is that the output, of the detector exceeds a threshold level. However, detection criteria can also be very complex. It is possible to design discrimination electronics, if the statistical natures of the signal and noise are known.

It should be noted here that since we are only interested in detecting transients, the photodetector is AC coupled to the electronics to allow detection against any constant intensity background.

Requirements for System Modeling and Design

The probability density function of the noise, at the input to the threshold detector, determines the threshold settings required to produce a given level of performance. Similarly, the power spectrum and time correlation function of the noise, at the input of the electronic filter, determines the passband of maximum signal to noise for a given signal.

The noise input to the filter is a sum of two components. One component, the dark noise of the detector and electronics, depends entirely on the type of detector and circuits used. The characteristics of this component may be estimated from knowledge of the detector material and configuration and a circuit diagram. They may also be measured for any completed device. If compensation for varying detector/circuit temperature is allowed, these characteristics are constants of the device.

The other component of the noise is caused by the random or pseudorandom fluctuations in the irradiance striking the optical filter. The portion of these varying light levels within the optical passband of the filter is passed on to the optics which images it on the detector. The detector converts it to variations in its output level. These unwanted fluctuations are passed on to the detection circuits. At this point, they combine with the dark noise to mask the weapons signals we wish to detect and produce false alarms. It is these fluctuations in light intensity that we must characterize and include if we are going to predict system performance.

There are two broad categories of phenomena which cause the fluctuations in light intensity. Variations in the amount of light reflected from the terrain cause changes in the background. Examples include sunglints from water, shadows of passing clouds, and waves on tall grass in the wind. Changes in the atmosphere also produce variations in the amount of light reaching the optics.

It is important to keep in mind the differences in these two phenomena. The total change in the terrain itself can be assumed to be the total of all the changes in all the individual components of the terrain. Variable reflections off terrain have definite magnitudes and spatial characteristics at their sources. For instance, a sun glint off water has a definite optical spectrum, temporal signature, intensity, and a limited spatial extent at the wave which creates it. These variations will decrease in magnitude and may change in other ways with the distance to the particular background disturbances causing them.

Fluctuations caused by turbulence in the atmosphere are not so easy to deal with. For instance, it is not at all clear how the intensity of such fluctuations change with distance to the limit of the field of view. Furthermore, since the light from the terrain must pass through the turbulent atmosphere, it is clear that the two types of noise are convolved not added.

The problem, then, is to measure the fluctuations of light induced by variable reflections off the terrain and a turbulent atmosphere. The statistical characteristics of background noise must be known for any set of system and background conditions.

We must be able to estimate such things as the probability density function of the background noise. Its power spectrum and time correlation function also must be known. In addition we must know how these characteristics change with various optical passbands and fields of view.

Some detection systems concepts are susceptible to variation in the shape of the image of the disturbance on the detector. Therefore, we must know the shape and size of the source of the light fluctuation. This information would be beneficial in distinguishing between terrain and atmosphere induced variations. We would also like to determine whether any classes of background transients exist which exhibit similar temporal or spectrum signatures. If such classes do exist it would be important to determine the statistical nature of their signatures so that possible discrimination schemes could be evolved.

Getting the Data
What Has Been Done

To date very little work has been done in this area. The most closely related field of active endeavor has been the effects of atmospheric turbulence on beams of light. Most of this work has been directed toward investigation and explanation of beam wander and non-uniform light intensities associated with the transmission of such beams through the atmosphere.

Perhaps the single most comprehensive work in this area is that of V. I. Tatarski of the Institute of Atmospheric Physics, Academy of Sciences of the USSR. In his book "Wave Propagation in a Turbulent Medium", Tatarski attempts a thorough theoretical explanation of the phenomenon of atmospheric scintillation. This is the phenomenon which causes the twinkling of the stars. He also attempts to establish agreement between his theoretical and experimental results.

Several others have made photographic studies of the shapes of shadow patterns produced by atmospheric scintillation on a beam of light. These studies show amorphous dark patterns which move across the beam with a velocity approximately equal to the component of wind velocity in that direction.

It is important to note here that such studies are of limited application to staring systems. Scintillation experiments measure the temporal and spatial variations imposed on an unmodulated beam due to transmission through the turbulent atmosphere. As such they are only applicable to a single point source of radiation at a known distance from the detector. Furthermore, the nature of these experiments is such that light entering the detector from points other than the experimental source

-504-

is rejected or ignored. Thus the concept of field of view does not apply.

A staring detection system may be considered as a mosaic detector located in the focal plane of collecting optics. Each elemental detector has a specific field of view. The total system field of view is the sum of the elemental fields of view. It is conceivable that both spatial and angular correlation functions are involved in characterizing background fluctuations.

Finally there is a very great difficulty in analyzing radiometric data of this nature. For although it is clearly possible to measure spatial, angular, and time correlation functions, probability density functions, and power spectral of signals from a mosaic of detector elements, there is no obvious way of separating the effects of terrain and those of the atmosphere once the data is taken. In fact, we do not now understand how two convolved effects can be "unconvolved" or separated.

This difficulty is also increased by the fact that the noise of the measuring system is included in such test results. Normally, to avoid such difficulties one merely makes the measuring equipment many times more sensitive than the detection equipment. However, in this case, the detection equipment is already pushing the theoretical limits of sensitivity for equipment of this type. Thus, the best measurements which can be made must either be made under extremely noisy background conditions or with a low background signal to system noise ratio.

Our Work to Date

A large part of our time and effort is spent measuring gun flashes. These measurements are made with radiometers which are quite similar to staring detection systems. When we noticed very noisy background conditions we recorded the resultant signals for 30 seconds or so. We then recorded an approximately equal length of data from the radiometers with the input aperatures completely blocked.

Our intention was to consider the differences in the power spectra of the two sets of data. The difference, it is hoped, will yield the power spectrum of the background (terrain plus atmosphere) component less measurement system noise. Thus, stationarity of the measurement system noise must be assumed. Presuming both background and system noise to be ergodic makes any measurement of the probability density function much simpler. However, it is not clear at this time what rational basis might be used to separate the effects of system noise from the probability distribution. This is complicated by the fact

that the AC coupling tends to bring the means of all such functions to zero.

Since we used single detection radiometers each with a different optical passband no spatial or angular data is available. Nor is any attempt being made to separate atmospheric effects from those of the terrain.

We are currently endeavoring to design equipment more nearly suited to these type measurements. This equipment will be based on an unusual detector developed for us by Minneapolis Honeywell. The detector is made of two layers of three bars each. The bars of the upper layer are at right angles to those of the lower layer. The layers are held together by a transparent insulating epoxy.

The two detector layers are made in such a way that the top layer responds to and absorbs one set of optical wavelengths while it is transparent to the wavelengths to which the bottom layer responds.

With this detector it should be possible to measure the angular correlation function over a 3 element spacing. This could be done in two optical passbands simultaneously and in the same field of view and from the same location. With such information we hope to begin trying to separate the effects of terrain from those of atmosphere. It seems knowledge of the angular size of the disturbance may allow us to separate those effects produced by wide shifts of meteorological conditions across the field of view. If detector arrays with more elements become available it should be possible to use visual sightings of events within the field of view to aid this separation effort. However, since the two components are convolved within any elemental field of view we know of no way to complete this process.

Summary of Needs   Thus a summary of our needs to complete the design of a successful experiment includes:

1. Some means of "unconvolving" two effects when some characteristics of one or the other are known or can be surmised.

2. Some means of eliminating the effects of measurement system noise from probability density functions, power spectra, etc of the measured data.

3. A means of measuring or inferring the physical shape of the image of the disturbances on the detector.

4. A theoretical treatment of the problem to allow overall guidance of the experimental effort.

Summary of Needs

Thus a summary of our needs to complete the design of a successful experiment includes:

1) Some means of "unconvolving" two effects when some characteristics of one or the other are known or can be surmised.

2) Some means of eliminating the effects of measurement system noise from probability density functions, power spectra, etc of the measured data.

3) A means of measuring or inferring the physical shape of the image of the disturbances on the detector.

4) A theoretical treatment of the problem to allow overall guidance of the experimental effort.

THE ANALYSIS OF A SUCCESS-FAILURE TIME SERIES WITH AN
APPRECIABLE NUMBER OF MISSING OBSERVATIONS

Robert P. Lee
Atmospheric Sciences Laboratory
White Sands Missile Range, New Mexico

ABSTRACT. The calculation of the power spectrum of a time series whose elements were either "success" or "failure" and where an appreciable number of the elements were missing is reported.

INTRODUCTION. This paper is the result of a request for a power spectrum analysis of a time series derived from surface meteorological data which had been collected hourly over about 19 years at White Sands Missile Range. Briefly, if certain wind, relative humidity, and time of day conditions were met and the visibility was less than a certain number of miles, the day was to be considered a success. A day which was not a success was labelled a failure unless no data was available for that day. Of the 6909 elements of the time series, 1641 were successes, 4784 were failures, and 484 were missing.

An examination of the data revealed that the missing elements were in general quite random, except for about four years near the middle of the time span during which very little Sunday data had been recorded. The Sunday data that was recorded showed what appeared to be a significantly higher percentage of pluses, indicating that the Sunday data during this period might not match the rest of the time series.

PROCEDURE. A plus one was assigned to a success, a minus one to a failure, and a zero was used as a key to the missing elements. Subtracting the mean from each of the data points and dividing by the standard deviation scaled the time series to unity variance with zero mean. The series now contained 1641 elements with a value of 1.71 and 4784 elements with a value of -.59, and as before 484 elements missing.

The first 197 autocovariance coefficients (1) were calculated from the equation

$$A_k = \frac{1}{P_k} \sum_{j=1}^{6909-k} X_j X_{j+k} \ , \ k=0, 1, 2, \ldots, 196$$

where P_k was the number of pair products present at lag k. (The value of 196 for the maximum lag was chosen, after several trials, because the exact reciprocal of 7, 14, and 28 days would be among the resulting frequencies.) The finite cosine transforms of the autocovariances were

Preceding page blank

smoothed using a three point filter (spectral window) having the weights
.25, .50, .25 producing the power spectrum shown in Figure 1. There are
two very obvious peaks at frequencies corresponding to periods of
(1) 7 days and (2) 3 $1/2$ days. The one labelled (3), almost indistin-
guishable in the noise, corresponds to a period of 2 $1/3$ days and is
in harmonic relationship with the first two. Since the time series had
unity variance, the sum of the power at all frequencies was unity,
and the power axis is so scaled. About 18% of the total power is in
the four lowest bands.

A suggestion was made that the missing data be filled in with pluses
and minuses drawn randomly from a population having the same proportion
as the original data. Since the background noise was already high, it was
assumed that the additional noise introduced by this procedure might not
be noticeable. The resulting spectrum is shown in Figure 2. Exactly
the same frequencies are still the predominant ones. In fact, the
spectrum is almost unchanged.

As a final test of the effect of the missing data and to eliminate
any possibility that bad data taken on Sunday had influenced the results,
all the Sunday data points for the entire 19 years were changed to
blanks. The resulting spectrum is given in Figure 3. Peak (1) has been
sharpened some, peak (2) is almost unchanged, and peak (3) is now in the
noise.

DISCUSSION. It is obvious, once it is pointed out, that it is possible
to wipe out points arbitrarily from a time series and still, using the
cosine transform of the autocovariances, obtain almost exactly the same
power spectrum. If the missing points are representative of the entire
set, the autocovariances, and therefore the power spectrum, will be
unchanged. In this case, removing the questionable data served to
raise the peak corresponding to a period of 7 days.

Two very valuable papers on the material covered in this paper have
been written by Dr. Richard H. Jones, now with the Department of Information
and Computer Sciences, University of Hawaii (2,3). They contain detailed
mathematical derivations of the effect of missing observations on the
variance of the spectral estimates, approximate degrees of freedom, etc.

It is to be expected that a large percentage of the energy present
would be in the very low frequencies corresponding to periods of one
year and greater, thus complicating the analysis at higher frequencies.

CONCLUSION. The power spectrum calculations show a predominant
amount of power at frequencies of one cycle per year and lower. From
meteorological experience, periods of 3 $1/2$ and 7 days are known to be
very reasonable. The results of this paper show that a time series
whose elements take only the two values "success" or "failure" is no bar

to the calculation of a reasonable power spectrum, even though an appreciable number of the observations are missing.

REFERENCES

1. Blackman, R. D. and J. W. Tukey, 1959. The Measurement of Power Spectra from the Point of View of Communications Engineering, Dover, New York.

2. Jones, R. H., 1962. "Spectral Analysis with Regular Missed Observations," The Annals of Mathematical Statistics, 33, 455-461.

3. Jones, R. H., 1971. "Spectrum Estimation with Missing Observations," Annals of the Institute of Statistical Mathematics, Tokyo. (In publication)
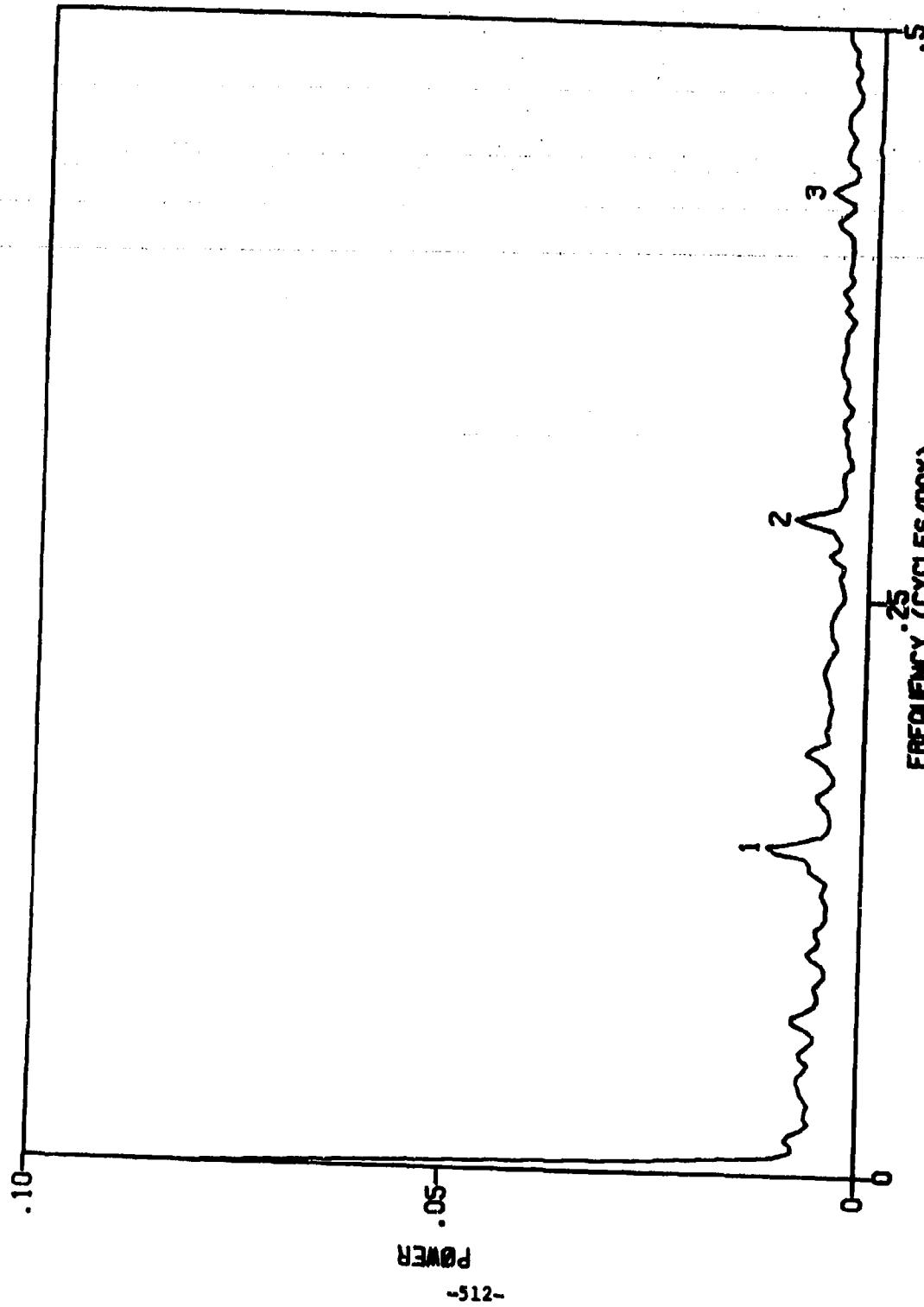
FIGURE 1 – POWER SPECTRUM OF DATA WITH MISSING OBSERVATIONS
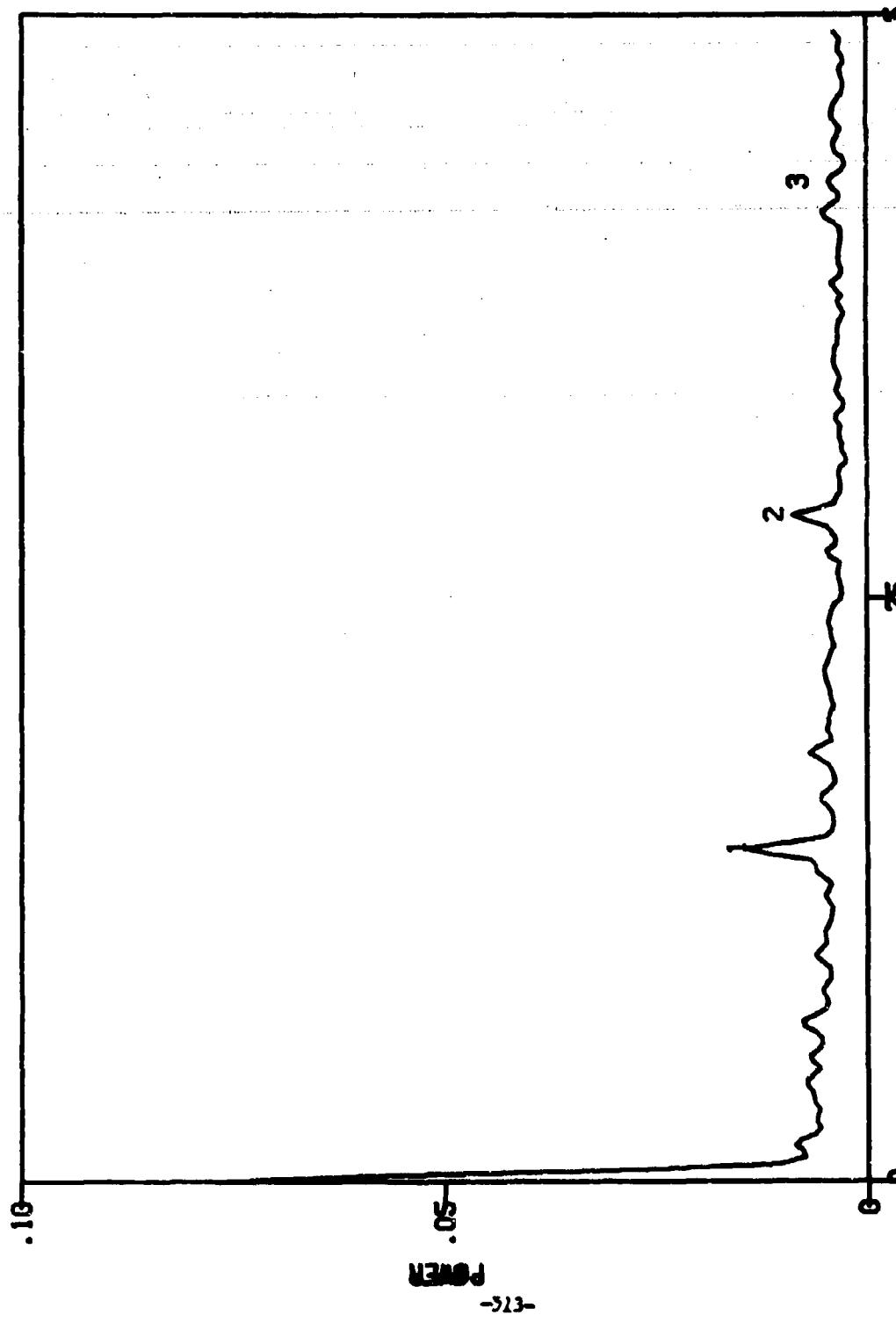
FIGURE 2 — POWER SPECTRUM OF DATA WITH MISSING OBSERVATIONS RANDOMLY FILLED

FIGURE 3 - POWER SPECTRUM OF DATA WITH ALL SUNDAY DATA ELIMINATED