PN-AAT-642
41834

ON-FARM AGRONOMIC TRIALS IN FARMING SYSTEMS RESEARCH AND EXTENSION

PETER E. HILDEBRAND AND FEDERICO POEY

Lynne Rienner Publishers, Inc.
Boulder, Colorado

Published in the United States of America by
Lynne Rienner Publishers, Inc.
948 North Street, Boulder, Colorado 80302

© 1985 by Lynne Rienner Publishers, Inc. All rights reserved

Library of Congress Cataloging in Publication Data

Hildebrand, Peter E.
On farm agronomic trials in farming systems research and extension.

Includes index.
1. Agriculture--Research--On-farm. 2. Field experiments. 3. Agricultural innovations--Research. 4. Agricultural systems--Research. 5. Agricultural extension work. 6. Appropriate technology--Research. I. Poey, Federico, 1933- . II. Title.
S540.053H55 1985 630'.72 84-27597
ISBN 0-931477-10-7 (lib. bdg.)
ISBN 0-931477-11-5 (pbk.)

Distributed outside of North and South America and Japan by Frances Pinter (Publishers) Ltd, 25 Floral Street, London WC2E 9DS England

Printed and bound in the United States of America

Contents

Analytical Techniques and Other Calculations viii
Preface ix
Collaborating Farmers xvi

I THE ROLE OF ON-FARM RESEARCH IN TECHNOLOGY DEVELOPMENT 1

Description of Farming Systems Research and Extension (FSR/E) 2
Purposes of On-Farm Research 5
Types of On-Farm Trials 6
Exploratory trials 6
Site-specific trials 6
Regional trials 7
Farmer-managed trials 7

II GENERAL CONSIDERATIONS RELATED TO ON-FARM TRIALS 9

On-Farm Research Practices 10
Researcher-farmer relations 10
Listening to and working with farmers 10
The nature of the relationship 11
On-farm experimental procedures 12
Location on the farm 12
Experimental designs 13
Field Data Management 14
Recording 14
Processing 15
Unusual values 15
Standardization of field data 16
Moisture and area 17
Maize shelling percentage 18
Missing plots 19
Missing plants 21
Analysis of covariance 22

III EXPLORATORY TRIALS 29

Superimposed Trials 30
The 2^n Factorial Trial 32
Interpretation of results 36
Comments on reducing the size of the trial 37
The "Plus" Trial 37
The "Minus" Trial 41

IV SITE-SPECIFIC TRIALS 45

Basic Considerations 46
Plot size 46
Variety evaluation 46
Crop associations 47
Plant nutrition 48
Plant protection 49
Analysis of Results 49

V RESEARCHER-MANAGED REGIONAL TRIALS: AGRONOMIC EVALUATION 57

Design and Methodology 58
Experimental designs 59
Number of replications 59
Number of treatments 59
Control treatments 60
Kinds of treatments 60
Selection of sites 62
Complexity 63
Replication over years 63
Combining Data over Sites 63
Combined analysis of variance 64

VI RESEARCHER-MANAGED REGIONAL TRIALS: SOCIOECONOMIC EVALUATION 73

Choice of Evaluation Criteria 74
Labor input as an evaluation criterion 75
Cash as an evaluation criterion 78
Response Surfaces 79
What they are 79
Representative forms 79
Economic analysis 81
Methods of Obtaining Response Surfaces 85
Simple regression--visiographic 85
Simple regression--least squares 92

Analysis of Response Surfaces 95
Coefficient of determination 97
Variance, standard error, and confidence intervals 98
Significance of the coefficients 99
Significance of the equation 100
Linear versus quadratic surface 102
A priori and *a posteriori* information 105
Multiple Regression--Visiographic 108

VII FARMER-MANAGED TRIALS 115

Design 116
Check plots 117
Number of sites 118
Analysis and Interpretation of Results 119
Farmers' evaluation 119
Passive evaluation 120
Active evaluation 121
Researchers' evaluation 125
Modified stability analysis 126
Distribution of confidence intervals 131
Response surfaces using modified stability analysis 142

VIII INITIATING AND MANAGING FSR/E PROGRAMS 149

Getting an On-Farm Research Program Started 150
Planning and Managing On-Farm Research 151
Defining Research Regions 153
Evaluating On-Farm Research 154
Regional Interinstitutional Cooperation 156

Bibliography 159
About the Book and Authors 162

Analytical Techniques and Other Calculations

Acceptability, index of 122
Coefficient of variation, CV 52
Confidence intervals, distribution of 131
Covariance, to standardize plant population 22
Data standardization for
  moisture and area 17
  missing plants 21
  missing plots 19
  maize shelling percentage 18
Factorial, 2^n 32
Mean separation
  difference between two means 40
  Duncan's new multiple range test 43
  Tukey's multiple range test 53
Modified stability analysis 126
  response surfaces 142
Profit maximization 82
Regression
  multiple-visiographic 108
    economic analysis 111
  simple-least squares 92
    coefficient of determination, R^2 97
    confidence band 100
    confidence interval around means 98
    linear 102, 128
    linear versus quadratic, ANOVA 104
    quadratic 92
    significance of coefficients, "t" test 98
    significance of equation, "F" test 100
    standard error, s 98
    variance, s^2 99
  simple-visiographic 85
Variance, analysis of
  after standardization for missing plots 21
  combined 64
  randomized complete blocks 51

Preface

The roots of the technological revolution in agriculture of the past century can be traced in the history of scientific discovery and the development of university research and training programs. Dramatic improvements in production per hectare and per farm, however, are a comparatively recent phenomenon. The first major breakthroughs occurred in the biological sciences with the development of hybrid maize in the 1930's, followed by the expanding use of complete fertilizers and improved weed and pest control technology following World War II. Scientific knowledge in the basic and applied agricultural sciences continues to advance at an accelerating rate and is the basis for confidence that food and fiber production can keep pace with growing world demand.

It is clear that not all farm households and family members have benefited equally from technological progress. Yields per hectare and per agricultural worker vary greatly among the regions of the world, among countries within regions, and among farms within each country. As a result, average farm incomes and the percentage of a nation's population engaged in agriculture also vary greatly among countries and regions. While the overall pattern of differences can be explained with reference to natural soil fertility, rainfall, product

demand, availability of inputs, and educational levels, these only partially explain the disappointing rate of progress in improving the productivity and income level of the smaller farms and those with relatively severe resource limitations.

Examining the history of research and extension (R&E) systems in the low- and middle-income countries of the world reveals four distinct phases. The first phase, which began prior to World War II, was marked by the construction of a limited number of research stations, either by public authorities or by international corporations interested in improved technology for commercial export crop production. Also, technical training programs were started, often outside colleges and universities. The focus of this first phase was on scientific research and the exploration of new crop production opportunities.

The second phase, dating from the postwar period, was characterized by a focus on rapid industrial development and by the rapid expansion of publicly supported community development and extension education programs for farm families and rural residents. The explicit assumption was that a backlog of technology available for adoption existed, so programs should focus on technology transfer and the motivation of target groups to accept change.

When this technology transfer strategy did not give the desired or expected adoption rates, a third phase was initiated in the late 1950's, with renewed emphasis on technology development and research institution building. Research programs emphasized genetic improvement, agronomic practices, and livestock management in an effort to identify "packages of practices" appropriate for small farms for which technology was not available. In this phase, the international research centers came into existence with their very strong, well-financed programs of genetic improvement for major food crop species. At the same time, cadres of scientists and agricultural specialists from developing countries were trained, leading to significant strengthening of their agro-biology research systems.

This third phase of agricultural research and extension system development is responsible for the marked growth in food production which has taken place in the past two decades. Total food production in developing

countries has been increasing more rapidly than in industrialized countries because of increased yields per hectare on some farms and because of a significant expansion in land under cultivation. Still, with higher population growth rates, the rate of increase in food availability per capita in developing countries has been less rapid than in developed countries. On the African continent, where agricultural research systems are less developed, per capita food availability has been declining.

Throughout this period, evidence began to accumulate that the supply of unused arable land was rapidly being exhausted and that existing research and technology transfer systems were still not meeting the needs of a majority of farm families who could be characterized as being on small-scale, limited-resource holdings. If adequate food and fiber production targets are to be met in the future, a new approach to applied research, technology development, and dissemination has to be implemented to generate the kinds of new technology acceptable to these limited-resource farmers.

Research and extension programs are now entering a fourth, client-participatory phase. The term "farming systems" was applied in the 1970's to several different activities that had common threads and similar purpose, but used different methods to pursue their goals. The common threads were:

1. A concern with small-scale, limited-resource family farmers who were reaping a disproportionately small share of the benefits of organized research, extension, and other developmental activities.

2. Recognition that a firsthand and thorough understanding of the farmers' situation is critical in increasing their productivity and helping to improve their welfare.

3. The use of scientists and technicians from more than one discipline as a means of understanding the farm as an entire system, rather than isolating components within the system.

Farming systems research and extension (FSR/E) is an approach to technology generation, evaluation, and delivery. It is applied, farmer-oriented, agro-biological research, supported by the socioeconomic sciences in a team effort that includes extension responsibilities. The

principal product is technology. The primary clients are farmers. Since FSR/E concerns technology generation, evaluation, and delivery, there are more agrobiological than socioeconomic scientists involved, and methodology emphasizes on-farm biological research as an integrated and critical portion of a sequence of activities.

In that context, this book addresses an important problem in agricultural technology innovation, namely that of technology development methodology. The problem is not new; it had been recognized explicitly before the rise in popularity of so-called farming systems research.

Traditionally, agricultural personnel have seldom made the important distinction between science and technology and between research and development. Counterparts in industry, with their well-known R and D designation, have long recognized the distinction. It is time that agricultural personnel did, and this book takes a step in that direction.

Technology is a synthesis, and technology development is synthesizing. Technology combines knowledge and other pieces of information into "something that works." Technology can be embodied in a machine, in a chemical product, in a seed, or in a cultural practice. Technology can be biological (seed), mechanical (machine), chemical (fertilizer), economic (policy), or intellectual (practice). A technology, to be useful, must serve without control over the other variables, and the wider the range of environments in which it can serve, the more valuable it is.

Agricultural technology is used in production systems. Thus, it must be tested in production systems, it must be adapted to production systems, and it must be integrated into these systems. An agricultural experiment station is not a production system. This simple truism has given rise to such terms as "farming systems research" and "on-farm research."

In the training of agricultural personnel, research methodology has always heavily emphasized science and a degree of control possible only in the laboratory or on the experiment station. The value of this activity and of science is not being challenged in this book. Most breakthroughs in agriculture have come from science, and the term "science-based" agriculture is an accurate one.

However, science and new knowledge is not enough. The new knowledge must be worked into a technology that in turn can be worked into a production system. Note the time span between discovery of the principles of hybridization and the use of hybrid maize in production. Note also all that had to happen in this period.

This book faces the technology issue squarely. It recognizes that farmers use technology, not science, even though the technology is science-based. It recognizes that most agricultural workers serving the technological needs of farmers deal more with technology than with science. It faces the need for an authentic methodology of technology development, to be used by those working in innovation without the luxury of laboratory and experiment station control. Technologists need their own technology. It is not enough to make do with the traditional research methodology, even with improvisations and ad hoc adaptations.

This book is an early step toward a methodology of technology development. It will be unfortunate, indeed, if it is the last word.

It is the on-farm biological research sequence that is the focus of this volume; its objective is to provide a practical guide to the design and analysis of on-farm agronomic experiments or field trials. Although FSR/E involves livestock components, as well as family and household components, this volume concentrates on the design and analysis of the agronomic components. The subject matter is organized following a logical sequence of the various types of on-farm trials: exploratory, site-specific, regional, and farmer-managed.

The book developed from a workshop on the design and analysis of on-farm trials in San Jose, Costa Rica, September 5-10, 1982. A number of experienced professionals participated in this workshop and an early version of the book was drafted. Original optimistic plans called for rapid editing of the draft and early publication. Subsequent revisions by the authors, an extensive search for empirical examples, and contributions by a number of other people eventually resulted in the present version.

Deciding what to include or add was a formidable task. The book reflects experience gained in the Farming Systems Support Project and has benefited from lessons

learned in training courses offered through the FSSP in North and South America, the Caribbean, and Africa. In its present form, the book does three things:

1) It presents the role and philosophy of on-farm research in FSR/E activities and describes a logical sequence for technology development.
2) It presents the most-used statistical procedures in simple, easy-to-follow steps. This is a service for technicians who are often isolated and would like to or must analyze their own data.
3) It presents new ideas and methods for analyzing agronomic data obtained without the effect of usual experiment station controlled conditions.

In an undertaking of this magnitude, many people are involved. Sponsors of the workshop were the Office of International Cooperation and Development (OICD), United States Department of Agriculture (USDA), who funded it; the University of Florida (International Programs and the Food and Resource Economics Department), who arranged it; and the Interamerican Institute for Cooperation in Agriculture (IICA), who hosted it. The facilities at IICA and the hospitality of its staff were greatly appreciated by the participants, who put forth an unusual effort themselves trying to draft a book in one week. The authors wish especially to acknowledge the efforts of the participants in San Jose, and those among them who reviewed the final draft and made useful suggestions.

Special recognition is made to the Instituto de Ciencia y Tecnologia Agrícolas (ICTA) of Guatemala for permission to use their data and to Juan Manuel Herrera, who, with the collaboration of former ICTA scientist Rene Velasquez, searched for and adapted many of the examples. Ramiro Ortiz, former Technical Director of ICTA, was especially helpful in making many suggestions regarding the statistical analyses and provided a great deal of help in editing several versions. Recognition is also due to CIMMYT, IRRI, and CATIE for the data furnished by them. Finally, particular gratitude is expressed to Jeannette Romero for her patience, understanding, and efficiency in preparing many drafts and the final version of the manuscript.

The workshop participants and their current affiliations in alphabetical order follow:

AGRIDEC, Guatemala	Ramiro Ortiz
AGRIDEC, USA	Federico Poey
CATIE, Costa Rica	Julio Henao M.
	Raúl Moreno
	Luis Navarro
Central Bank of Ceylon, Sri Lanka	Anila Wijesinha
CONACYT, Ecuador	Franklin Arboleda
ICTA, Guatemala	Juan M. Herrera
IDRC(CIID), Colombia	Nicolás Mateo
IICA, Peru	Antonio M. Pinchinat
IICA, Costa Rica	Rufo Bazán
	Victor Quiroga
	Jorge Soria
	Karel Vohnout
	Quentin M. West
INTA, Argentina	Guillermo Joandet
Lynne Rienner Publishers, USA	Lynne Rienner
North Carolina State University, USA	Larry A. Nelson
OICD/USDA, USA	Donald Ferguson
University of Florida, USA	Peter E. Hildebrand
	Robert K. Waugh

Others who reviewed a final draft and made useful suggestions were Louise Fresco, Wageningen Agricultural University, The Netherlands; Tom Stillwell, Michigan State University; and Steve Kearl, Ken McDermott, Chris Andrew, and others of the Farming Systems Support Project at the University of Florida. We gratefully acknowledge their contributions. The authors, of course, accept all responsibility for any errors of omission or comission that may still exist, as well as for the final content of the book.

Peter E. Hildebrand
Federico Poey
Gainesville, Florida
October 1984

Collaborating Farmers

Collaborating farmers are thinking human beings. They know the characteristics of their land better than the technician; they have worked it for many years. They are proud, have self-esteem. They are pragmatic: if they invest in their land they expect tangible results, so much the better if obtained in the short-run. They keep using "wrong" practices which they believe in although it may be difficult to determine the source of their beliefs. They do not like to waste time. They are appreciative and hospitable. They are sure to obtain better results than the researchers if both use limited resources. They perform various activities in the cropping cycle according to the phases of the moon. To be polite, they may indicate understanding what they have not understood. They are interested, more than anything else, in their current crop or crops. They want to show us how much they know, and they know a lot. They like to ride in cars. They don't bother talking about things that do not relate to agriculture. Their income and education are low.

For these and other reasons:

1. Let's try not to make mistakes either in or beyond their presence.
2. We must not lie to them.
3. Let us be punctual.
4. Let us respect their point of view.
5. Let us not take advantage of their oversights or lack of knowledge.
6. Let us talk less and work more.
7. Let us make them participants in our activities and help them understand what we are doing.
8. We should repeat what they have not understood until they really do.
9. Let us treat them as friends.
10. They have a great deal to teach. Let us learn from them.

Lee Roy Gillespie
Translated and adapted
from Noticta
November 1978
ICTA, Guatemala

I
The Role of On-Farm Research in Technology Development

The client-participatory approach to the generation, evaluation, and dissemination of technology developed in recent years involves a sequence of activities in which the clients (in this case, usually small-scale, limited-resource family farmers) are involved in most of the steps. This sequential procedure, known as Farming Systems Research and Extension (FSR/E), is flexible and adaptable to different conditions encountered in the field and in the institutions involved. It is iterative in the sense that new information is used immediately and is also fed back into the sequence to improve earlier stages being repeated in another cycle.

Initial activities involve a characterization of the farming systems in an area, through discussions with the farmers themselves and through the tentative partitioning of the systems into homogeneous groups, or recommendation domains, which become the basis for making specific technology recommendations. A major portion of the biological research conducted in an area to help solve problems encountered there is carried out on farms with the participation of the farmers. Eventually the farmers are asked to manage simple trials themselves in order to assess the acceptability of the technology when it is completely under their control.

DESCRIPTION OF FARMING SYSTEMS RESEARCH AND EXTENSION (FSR/E)

Although FSR/E is flexible enough to adjust to the agricultural and institutional conditions existent in different countries and cultural settings, it will usually involve a sequence of steps similar to the following within a previously determined geographical or political region.

1. Initial characterization and analysis of existing farming systems through close consultation with farmers
   a. First estimation of problems and constraints
   b. Tentative partitioning into homogeneous farming systems, or recommendation domains

2. Planning and design of first-phase work
   a. Biological research
   b. Continuing agro-socioeconomic characterization

3. Selection, generation, and evaluation of technologies
   a. Commodity and discipline research on experiment stations and laboratories
   b. Researcher-managed on-farm trials with farmer participation
   (i) Exploratory trials
   (ii) Site-specific trials
   (iii) Regional trials
   c. Farmer-managed trials
   (i) Individual evaluation of acceptability by the farmers themselves
   (ii) Refined partitioning of recommendation domains by researchers
   (iii) Initiation of technology transfer activities

4. Collection and analysis of data
   a. Agro-technical data from on-farm and on-station trials
   b. Economic records of farm enterprises from farmers
   c. Additional agro-socio-cultural and political information from farmers and other area residents

5. Frequently programmed multidisciplinary re-evaluation of research activities and information to do the following:
   a. Redefine partitioning of recommendation domains
   b. Make recommendations of acceptable technology for dissemination into specified recommendation domains
   c. Introduce feedback into the sequential process
   d. Serve as a basis for planning future work

6. Promotion of acceptable technology throughout appropriate recommendation domain(s)

In many ways, this sequence parallels what farmers have always done. Farmers manage a complex set of biological processes which transform the resources at their disposal into useful products, either for home consumption or for sale or trade. The choice of crop and livestock enterprises and the methods and timing of cultivation, husbandry, and harvesting are determined not only by physical and biological constraints, but also by economic and sociopolitical factors which make up the larger milieu within which the farmers operate. Conceptually, there are many sets of choices and outcomes which would have direct consequences on the welfare of farm families.

Within this complex milieu, through a process of trial and error and over a number of seasons or generations, farmers move toward appropriate technologies and allocations of resources, given their specific objectives. While the choices available to each farmer are different, those with similar sets of resources and constraints tend to make similar choices as to crops, livestock, and management practices. Those who have responded in similar ways can be grouped together into homogeneous farming systems. The current technology they are using, which has evolved over a long period of time, will be similar within these similar groups.

FSR/E brings scientific method and additional expertise to bear on this process of problem identification and technology generation. Teams of scientists from different disciplines, working with farmers, can speed up the process and make it more efficient in responding to a rapidly changing world. The on-farm trials (3b and 3c, above), represent a sequence

designed to assure the advantages and acceptability of new technology by the collaborating farmers. Materials and methods that move through this evaluation phase come from experiment stations and other sources. Depending upon the nature of the technology being evaluated, it is usually possible to initiate on-farm activities with site-specific or regional trials.

On-farm research is not a substitute for experiment station research--it is a means of providing much wider exposure to station results, both with respect to environment and to potential users. It is also a means of conveying to station researchers any problems that require experiment station facilities for solution. That is, on-farm research provides an opportunity for station researchers to expose their results to a much wider range of environmental conditions. On-farm research also provides an opportunity for more and smoother interaction between extension personnel and the research procedure.

In moving through the sequence from experiment station results to extension and farm production, the complexity of the trials (number of treatments and replications) at each location diminishes as plot size and number of locations increase. In this sequence, the extent of farmer management of the trials increases, and the need for researcher management decreases (making possible the larger number of locations). Concomitantly, the capability and need to control sources of variation decreases, while the need and possibility of measuring the sources of variation increases. As the above changes occur, biological precision and discrimination among variables decreases, while the ability to test socioeconomic interactions under farmers' conditions increases. All of the above changes increase the number of farmers involved in formalized research and increase the direct investment farmers make in that research. Finally, as the number of farmers increases, the potential interaction of extension with research is enhanced.

PURPOSES OF ON-FARM RESEARCH

On-farm research in this context can become a focal point for developing a technological system to serve farmers, opening up several new possibilities for improving the effectiveness of research:

1. On-farm research can serve as a linkage for ongoing research and extension and the improvement of both.
2. It can make component research more purposeful. It serves as a basis for evaluating the output from discipline and commodity research, because it can function to integrate the results from that research.
3. It can serve as a basis for the orientation of component, commodity, and discipline research and the selection of priorities.
4. It can make research more comprehensible, and therefore more attractive to decision makers.
5. It can furnish information and introduce checks and balances (evaluations) that can improve research and extension management.
6. It can be a hands-on experience to improve the effectiveness and the image of research and extension workers, typically viewed by farmers as inhabitants of ivory towers who do not understand the reality of farming.
7. It can add to biological research, making it more effective by evaluating responses when the non-experimental variables, including management, are allowed to fluctuate within the farmers' normal conditions of production. The conventional research system gives an estimate of what would happen if farmers were to control variables as the researcher does. It does not, however, furnish an estimate of results if farmers were to actually use the new technology. Both estimates are important, but without on-farm research, the latter is missing.
8. The entire sequence can be considered as a learning process for researchers, extension personnel and farmers. It helps to refine both technology and the definition of the recommendation domain(s) for which specific technology is appropriate.

TYPES OF ON-FARM TRIALS

Types and numbers of trials are planned for each recommendation domain previously identified by an initial characterization of a region. The nature of the problems, the availability of personnel, and budget considerations all influence this allocation. Except for exploratory trials, which can be used at any time to learn about unknown situations, the other on-farm trials are sequential, with specific purposes at each stage.

Exploratory Trials

Exploratory trials are used when little is known about an area or about possible effects in an area of a specific type of technology. They can be considered as complementary to, or part of, characterization and usually precede site-specific or regional trials. These trials normally provide more qualitative than quantitative information about several factors. Frequently, two levels of each factor are included and few replications are used. The most common designs are the 2^n factorial and plus or minus trials. Exploratory trials can sometimes be superimposed on farmers' fields without the necessity of special preparation of the experimental area.

Site-Specific Trials

These are similar in design to on-station trials, but usually fewer treatments are involved. Perhaps as many as 20 to 25 treatments can be included, although this is not recommended unless a more complex type of design (e.g., lattice or Latin square) is used to keep the experimental error at an acceptable level. Because of the requirement for intensive researcher management, few of these trials are normally conducted. The most common design is randomized complete blocks with four replications. Analysis of variance (ANOVA) can include site as a source of variation, and combined analyses can be performed.

Regional Trials

Regional trials are amenable to agronomic and agro-socioeconomic analysis. They are designed to expose the best treatments from site-specific trials to a much wider range of environments within a recommendation domain. Perhaps six of the best treatments are included, and five to ten sites can be utilized. A recommended design is randomized complete blocks with two to four replications per site. ANOVA, regression, or modified stability analysis (see Chapter VII) can be utilized. Combined analysis with site as a source of variation can be used in ANOVA.

Farmer-Managed Trials

These trials provide the opportunity for the farmers themselves to manage and evaluate the one or two most promising treatments from regional trials. Large plots with no replications are used. The purpose is for the farmers to be able to compare the treatments with their own practices, so one plot with these practices can be included in the design. In practice, this check plot serves the researchers more than the farmers, because the farmers will be able to evaluate results based on their own fields. If researchers wish to measure results of the farmers' own practices, they can also sample from the farmers' fields. However, agronomic and economic records of the farmers' practices must be kept to provide the necessary information. It is desirable to have at least 30 farmers in these trials in a recommendation domain. Larger numbers improve the precision of the conclusion, but smaller numbers can still provide useful information.

The remainder of this book deals with considerations related to on-farm trials; the different kinds of on-farm trials are discussed in separate chapters. Stressed throughout is the concept that each kind of trial is part of a sequence through which technology passes as it is being designed, evaluated, and disseminated. None of the steps in this sequence is sufficient in and of itself, and all, taken together, depend on other on-farm research not covered in the book. Some of these are preliminary or

special surveys, farm production records, and other formal and informal contacts with the farmers and other residents in the area.

II
General Considerations Related to On-Farm Trials

Management practices and field conditions on most farms differ from those found on experiment stations. These differences need to be considered in any strategy to obtain meaningful experimental data from on-farm trials. On-farm trials are not meant to try to simulate experiment station conditions in farmers' fields. Rather, they are designed to help detect differences under typical farmer management practices and environmental conditions.

On-farm research is characterized by farmers' participation on their own land. This participation varies according to the nature of the experiments. In exploratory and site-specific trials, it is limited to providing the land and some or all of the inputs. At this stage, farmer participation in information gathering and decision making is secondary to that of the researcher who controls the trials. In regional trials farmer participation is greater, contributing heavily to the interpretation of results and eventual recommendations. Finally, farmer-managed trials are conducted by the farmer, while the researcher becomes the collaborator.

Researcher-farmer relations, location of trials on the farm, on-farm experimental designs, and field data management, including recording, processing, and standardization, are a few of the many facets that need to be viewed from a proper perspective when doing research in

farmers' fields with their active participation.

ON-FARM RESEARCH PRACTICES

Researcher-Farmer Relations

When conducting research on farms, researchers are intruding upon the farmers' land and taking their valuable time. The research may be using other of the farmers' scarce resources. Because of this, it is well for the researchers to act always in the best interest of the farmers, treating them as equals in the research process and considering them as desirable, not just necessary, components in the technology generation, evaluation, and dissemination procedure. Farmers understand experimentation and are willing to participate if they feel they will possibly benefit from it, and if they understand what is happening. It is of utmost importance for researchers to explain fully why they are there, what they would like to do, what is going to be required of the farmers, and what the farmers can expect from the results. It is most important to explain why it will be of value and of interest to the farmers to be participants in the undertaking.

Listening to and working with farmers

From the very first contact made with farmers in the initial survey, or in looking for collaborators for on-farm trials or enterprise records, it is extremely important that the researchers begin by listening to and working with the farmers. Farmers resent being told by "government people" that they are doing things wrong, and that the "outsiders" know how the farmers should do it better. If the researchers convey this attitude to the farmers from the beginning, the relationship will get off to a slow start, if it gets started at all.

Care must be exercised by the researchers to ascertain which of the household members are the decision makers and to talk with those who are responsible for specific crops. A wife may know little about her husband's cotton crop; he may know little about her cassava or peanut crop.

The nature of the relationship

Farmers should be aware from the beginning exactly what to expect from the relationship. Above all, they must be informed that the work is research, from which both researcher and farmer will learn, and not a demonstration designed to show how much better the researchers can do what the farmers are already doing. (In most cases, the farmers know how to do it better, but they cannot afford to.) Farmers must be aware of who will be expected to provide what, who will take what risks, who will get what product. It is critical that farmers understand the timing of the various activities and whether it is to be at their initiative or at the initiative of the researchers. For example, in a yellow maize area, if some white varieties are to be used, the farmers should know if they can expect some yellow maize in return for the white maize they will not want, or if they should just expect to lose that which was produced. They should also agree to include white maize and understand why it should be included. They must know who should provide the fertilizer, if it is to be used, and when it must be available; who is going to harvest, when and how.

Farmers understand risk and are willing to (or are forced to) accept it as a normal part of their production environment. If an experiment is lost because of normal environmental conditions, farmers will understand it and will not be concerned about compensation (although they would probably accept it if offered). In order to avoid paternalism in the research process, it is better not to consider compensation for these cases. If, on the other hand, certain treatments are lost because they were poorly thought-out or obviously not adapted to the production environment of the farmers, the farmers can be expected to think compensation is warranted unless they were well advised beforehand of this eventuality. In this case, payment in kind, of the quantity and quality that otherwise would have been produced, is probably indicated. It is better, of course to avoid the situation by having well-thought-out, simple interventions and adequate farmer involvement in the design of the trial.

Farmers must understand the importance of the trial to the researchers. The risk of not completing on-farm

trials is higher than with experiment station trials, because much depends on the cooperation of the farmers. There are many examples of "lost" on-farm trials due to decisions made by the farmers without consultation with the researchers. An increase in the market price of the product might cause a decision for an early harvest of part or all of the trial. A new variety or crop that is considered especially attractive might promote harvest by farmers or their neighbors before the final data are recorded. Under some circumstances, preliminary results satisfy the curiosity of the farmers and they lose interest before the trial is completed. When trials involve more than one cycle of production, or when it is necessary to evaluate a rotation of crops, the risk of not completing an on-farm experiment increases.

Farmers who do not fully comprehend the nature of the trial may enter into competition with researchers. For example, a check treatment that is meant to simulate the farmers' practices and is to be conducted by the farmers may receive special care because the farmers know how to do it better and want to prove this to the researchers. On a small plot, they can afford to do it even if they cannot do it on their own fields. Or, the farmers may not understand fully that they are supposed to manage the plot exactly the way they do their own fields, so they wait for the visits of the researchers before they carry out practices that they normally do earlier on their own land. In either case, errors are created in measuring the farmers' level of production.

Finally, periodic review of all aspects of the trial, along with frequent conversations between the researchers and the farmers concerning the progress being observed, is critical to fruitful on-farm research.

On-Farm Experimental Procedures

Location on the farm

Homogeneous or uniform experimental areas are the rule rather than the exception on experiment stations. The opposite is true on farms. Nevertheless, researchers can reduce experimental error by following a few common-sense rules. For example, it is never wise to locate a research area adjacent to a habitation unless that is the

environment in which the crop in the trial is going to be planted normally. Likewise, paths, canals, large trees, and other conditions which are not normally part of the environment should be avoided. If the crop is usually planted in these special environments, of course, it is appropriate to locate the experimental area in them.

Experimental designs

Conducting field trials on farms does not mean that scientific methods can be overlooked. The same basic methods are used as for any other research. The experimental design or arrangement depends on the results of the preliminary reconnaissance of the region, the variables to be measured or controlled, environmental variability, and the specific objectives of the trial. It should be stressed that a good design is essential for trials conducted on farms, as it is often the key to helping farmers retain their beliefs and confidence in the research institution and in the researchers themselves.

Using specific designs simply because they were used elsewhere in similar experiments is discouraged. Whether or not to use blocks, how many replications to make, plot size, and other design considerations will depend on the particular problem in the particular location. The number of controlled variables and the amount of data collected should not be more than necessary for attaining trial objectives. There is a natural tendency to record as much information as possible, on the assumption that it might help explain findings that may emerge upon completion of the trials, or simply that it might be "interesting." But experimental information, particularly at the farmer's field level, is costly to obtain, and it is preferable to limit data recording only to those data which are useful.

Another consideration in planning a field trial is to specify who will conduct it. The degree of complexity will depend largely on who will be in charge of implementation. Implementation may be assigned to personnel working on an experiment station, or to a multidisciplinary field team that planned the trials, or perhaps to technical assistants trained for this purpose. Extension agents or farmers may also be involved in the management of the trials. Farmers' participation, in particular, should be carefully defined in order to make

their contribution as effective and valid as possible. It should be remembered that they also participate in reaching conclusions and recommendations from the trials.

The most common experimental design used in on-farm research is probably randomized complete blocks, usually with four replications. Split-plot arrangements are not encouraged, but may be necessary when ecological conditions or the nature of the variables prevent a complete randomization of plots, as, for example, when comparing fertilizer levels with and without irrigation. The fertilizer levels are randomized within larger blocks that are either irrigated or rainfed. Another example would be to minimize cultivar border effects by randomizing plant populations within larger blocks that are cultivars.

FIELD DATA MANAGEMENT

Recording

By its very nature, the information recorded in on-farm trials must be less than in trials conducted on experiment stations. The minimum necessary information should be taken. Since experiments on farmers' fields cannot receive the same day-to-day attention as station experiments, it is advisable to increase the number of locations rather than become involved in data collection with too much detail in fewer locations. Those in charge of on-farm trials must make every possible effort to reduce to a minimum the time between completion of recording data and the issuing of recommendations. Never forget that farmers develop high expectations when something is done on their farms, and their curiosity must be satisfied as quickly as possible if their support and assistance are to continue.

It is important to decide how records are going to be managed before experiments are conducted. Developing standard procedures of data recording helps speed processing and analysis, and contributes in turn to faster conclusions and earlier recommendations. It also makes information more reliable and easier to file and retrieve. The availability or absence of electronic data-processing facilities should be a prime consideration in planning recording techniques.

Recording sheets should meet the following requirements:

a. Field data should be usable directly for processing; eliminating transcriptions saves time, cuts the costs of trials, and eliminates one source of error.
b. There should always be at least one original and one copy of records.
c. Records should be easy to read; this implies not too many records per sheet.
d. Sheet size should be such that it can easily be handled in the field, as well as filed in standard files.

Processing

Before data are analyzed, they must be inspected for irregularities, processed, and standardized.

Unusual values

Before any data analysis is attempted, the patterns of variation in the data should be studied. Attention should be given to numbers which appear to be unusual (called outliers), and those which are missing (called missing plots). An attempt should also be made to see if the variation is homogeneous throughout the data set. The range is a useful device for this. An example of the use of the range, to look for non-homogeneity of variation, is illustrated with field data from a trial with six treatments and three replications:

	BLOCK I	BLOCK II	BLOCK III
Treatment 1	40	60	80
Treatment 2	30	55	120
Treatment 3	20	70	92
Treatment 4	20	42	60
Treatment 5	40	58	80
Treatment 6	50	68	92

The range in yield for each of the treatments is obtained by finding the difference between the highest and lowest values in each treatment:

Treatment 1	80-40 = 40
Treatment 2	120-30 = 90
Treatment 3	92-20 = 72
Treatment 4	60-20 = 40
Treatment 5	80-40 = 40
Treatment 6	92-50 = 42

Treatments 2 and 3 have very large ranges, so the data should be inspected further to find out why. Upon inspection, the value 120 for Block III, Treatment 2 seems too high, and the value 20 for Block I, Treatment 3 seems too low. The researcher should look for specific physical reasons why these numbers are unusual. Sometimes they can be traced to copying or typographical errors or to some unusual situation that occurred in a plot but did not affect the other plots. If a specific reason not associated with the experiment can be found, the numbers may be replaced by new values obtained by checking original field records, using missing plot formulas, covariance, or other suitable methods.

Standardization of field data

The field information taken directly from the experimental plots (raw data) can seldom be utilized as such for statistical analyses. Depending on the type of crop, time of harvest, part of the plant of interest, and many other factors, it is usually necessary to make some numerical transformations that will provide more reliable interpretations of the data. A common correction is made when comparing yields of maize varieties with different rates of maturity; if grain moisture is not standardized to a uniform content, the excess moisture in the grain of late-maturing varieties will cause an upward bias for those varieties if direct plot weights are interpreted. Also, plot size needs to be transformed in order to produce more meaningful values. For example, it is better to interpret tons, or kg/ha, of grain at a constant moisture, than to consider just kilograms or grams per plot, with no reference to the moisture content or plot size. The correction procedures for these and other factors are illustrated with field data.

Moisture and area. Table II-1 describes yield and moisture content at time of harvest and the converted yield in kg/ha at 14% moisture. There were three blocks containing plots of 50 m^2 net area.

TABLE II-1. Field data and standardized yield of wheat from a phosphorus experiment in Guatemala.

Block	P. level	kg/plot	Field Moisture %	kg/ha at 14% moisture
01	0	3.08	20	569
01	40	3.68	21	676
02	0	6.52	22.0	1183
02	40	7.44	20.0	1384
03	0	6.25	19.8	1166
03	40	6.28	20.0	1168

Source: ICTA

Correction Factor for Area (CFA) to convert weight to kg/ha is calculated as follows when plot size is measured in square meters (m^2):

CFA = 10,000 / net plot size

The Correction Factor for Moisture (CFM) to convert weight to a constant moisture content is estimated as follows:

CFM = (100 - % hvst. moist.) / (100 - % constant moist.)

As an example, in Table II-1, for the first row the calculations are:

CFA = 10,000 / 50
= 200

CFM = (100 - 20.5) / (100 - 14)
= 0.924

Then, for the first row, kg/ha at 14% moisture is

3.08 x 200 x 0.924 = 569

PRACTICAL FIELD ADVICE
The farmer estimates yield without correction for moisture. If treatment differences are so small that moisture adjustments must be made, the farmer will not be able to detect them.

Maize shelling percentage. In maize, field weight is normally reported as kilograms of ears per plot. Converting these values to grain weight is necessary when comparing varieties that differ in the ratio of grain-to-cob weight.

PRACTICAL FIELD ADVICE
The farmer will often express this difference by saying the variety "does not yield as much," meaning a net or basketful or other standard farm measure of ear volume does not "yield" as much grain.

The Correction Factor for Shelling (CFS) percentage can be obtained from the grain and ear weight of a random sample of ears, as follows:

CFS = kg of shelled grain / kg of ear corn

For example, if 20 ears weigh 4.1 kg and the shelled grain weighs 3.3 kg, then

CFS = 3.3 / 4.1
= 0.805

This correction factor is then multiplied by the total ear weight of each plot of the same variety.

Missing plots. Frequently in on-farm research, animal intervention or other unusual occurrences can ruin one or more plots in a trial. A decision must be made by the researcher on how to adjust the trial to account for these missing data. There are several ways to do this: generate a value for the affected plot, drop the block or replication from the analysis, or analyze all remaining plots as if they were a fully randomized design with unequal numbers of replications. Still another alternative, if the plot is not completely destroyed, is to harvest the parts of the plot that are undamaged and proceed as for missing plants (see next section).

If fewer than four replications were used in the original design, dropping an entire replication is a fairly drastic measure, and other alternatives should be considered. If only one or two plots were affected and there were several treatments in the trial, then generation of estimated values would be the best alternative. If regression rather than analysis of variance is to be used to analyze the data, then a missing plot is less of a problem and may be omitted without significantly affecting the analysis.

Standard statistical texts recommend a procedure not too complicated for field use, if only one, or at most two, plots are missing. For a randomized complete block design a single missing plot value can be estimated by the following equation:

$$Y = (bB + tT - G) / (b - 1)(t - 1)$$

where b and t are the numbers of blocks and treatments, respectively, B and T are totals of observed plot values in the block(s) and treatment(s) containing the missing information, and G is the grand total of all observed plot values. The calculated or estimated value Y is entered into the data where the plot was missing. Analysis of variance is performed as usual, except that one degree of freedom is subtracted from total degrees of freedom (and therefore error degrees of freedom will also have one less than if the plot value had not been missing). Treatment sum of squares will have to be reduced by an amount equal to:

$$[B - Y(t - 1)]^2 / t(t - 1)$$

where Y comes from the previous equation. An example follows:

Treat-ments	BLOCKS I	II	III	T_i
D	30	26	( )	56
E	35	27	33	95
F	43	35	45	123
B_i	108	88	78	274

$$Y = [3(78) + 3(56) - 274] / (3 - 1)(3 - 1)$$
$$= 32$$

The value 32 is used in block III for treatment D. For the analysis of variance, ANOVA (described in detail later in this chapter), total degrees of freedom (d.f.) is $(rt - 1) - 1 = 7$ and error degrees of freedom will be $[(r - 1)(t - 1)] - 1 = 3$. Treatment sum of squares when calculated is reduced by $[78 - 32(3 - 1)]^2 / 3(3 - 1) = 32.67$. Then treatment mean square and the F value can be calculated.

Sum of squares calculated with the estimate of missing plot value:

Source	d.f.	Sum of squares
Blocks(r-1)	2	98.67
Treatments(t-1)	2	228.67
Error	4	10.67
Total rt-1	8	338.00

Note: r = Number of replications or blocks
t = Number of treatments
d.f.= degrees of freedom

ANOVA as adjusted:

Source	d.f.	Sum of squares	Mean square
Blocks	2	98.67	
Treatments	2	228.67-32.67= 196.00	98.00
Error	4-1 = 3	10.67	3.56
Total	8-1 = 7	305.34	

F_C = 98.00 / 3.56
= 27.53

CV = 5.55%

Note: F_C = Calculated F value
= Mean square of treatment/error mean square.

The calculated F value, F_C, is larger than the 5% F value found in a table for 2 and 3 degrees of freedom (9.55), so there is a significant difference among treatments at the 5% level based on the adjusted ANOVA. Had the ANOVA been performed with the non-adjusted figures it would have indicated significance at the 1% level.

Missing plants. Other common corrections include an adjustment for plant population when this factor is affected by an outside influence that is not part of natural environmental conditions. This is the case, for example, when animals (or persons) interfere in an experiment, removing or damaging plants. Since population correction procedures tend to favor (increase) treatment values, these should be made only when truly justified. Judgement should be exercised in applying these correction factors, because generally, in on-farm research, the differences sought among treatments is much larger than effects from usual plant population variability. This judgement should take into consideration the variability that can be attributed to normal environmental or local conditions. For example, if germination is affected by normal environmental conditions, correction for stand should not be practiced. This is the case in some parts of the highlands of Guatemala where maize is planted very deep two months before the rainy season starts. If an experimental variety does not have the ability to withstand that condition, correcting for plant population would be a mistake.

PRACTICAL FIELD ADVICE
A practical method of adjusting plot yield values for missing plants is to harvest only those plants that have full competition, then calculate the appropriate area for that number of plants and adjust for normal plot size. This can work for crops such as maize or for many vegetables, where individual plants are easily distinguished and harvested and for which a specified number of plants is expected in a full plot. For example, if 25 plants should be in a plot and only 18 are found with complete competition, then the yield for these 18 can be increased by multiplying by a factor of 25/18 to get the estimated yield for the full plot, had it not been damaged by outside causes. An alternative method for crops which are not so easily separated into individual plants, or for which plant numbers are not calculated for the plot (such as wheat), is to locate small areas in the plot which have not been damaged and harvest them, leaving an unharvested boundary. Then the total area harvested can be adjusted to the size of the normal plot and yield adjusted accordingly. In both of these cases, the implicit assumption is that the parts of the plot harvested were representative of the whole plot, a possible source of increased experimental error. However, this reduces the calculations which are necessary for more sophisticated adjustments.

If correction needs to be made, by no means should a direct relationship of the average weight of all remaining plants be used as the estimate. Those plants, individually, would have higher than normal production because of a lack of nearby competitive plants.

Analysis of covariance. A more complete and accepted method to standardize plant population is through the analysis of covariance of plot weights and number of plants reported. The following explanation and example illustrate the use of this method.

The analysis of covariance is a statistical method that allows valid treatment comparisons using observations

of one variable (yield), after the effect of a possible disturbing variable (incomplete plant stands) has been removed. The application of covariance must be justified in the sense that the correction is being made because of an uncontrolled environmental condition that does not affect all the observations (plots) in the trial in a homogeneously constant fashion. The example to be presented in this section is a correction of yield in a trial where animals ate plants in different plots, more in some than in others. This exogeneous effect is not related to the ability of the cultivars under evaluation to withstand adverse environments. However, if grazing is part of the environment in which the crop will be produced, corrections should not be made.

The example presented is an evaluation of six new maize hybrids in a randomized complete block design with four replications.

Step 1. Table II-2 shows the of number of plants (X) and kilograms per plot (Y) for each cultivar.

TABLE II-2. Number of plants (X) and kilograms per plot (Y) of six maize hybrids.

Treat-	BLOCKS									
ment	I		II		III		IV		TOTAL	
No.	X	Y	X	Y	X	Y	X	Y	X	Y
1	60	3.42	59	5.25	62	4.52	60	6.24	241	19.43
2	47	2.87	60	3.97	61	3.12	51	4.82	219	14.78
3	51	4.21	35	3.29	60	5.58	40	3.82	186	16.90
4	58	2.58	32	1.93	62	3.07	50	4.05	202	11.63
5	62	3.28	54	4.13	60	4.05	62	5.38	238	16.84
6	40	1.98	62	5.02	61	3.89	61	5.20	224	16.09
TOTALS	318	18.34	302	23.59	366	24.23	324	29.51	1310	95.67

Source: ICTA, Guatemala

Step 2. Perform an analysis of variance (ANOVA) for number of plants (X) to get the sum of squares (Σx^2) values:

$$
\begin{aligned}
(CF) &= \text{Correction Factor} \\
&= (1310)^2 / 24 \\
&= 71504.17
\end{aligned}
$$

$$
\begin{aligned}
SST &= \text{Treatment sum of squares} \\
&= \{[(241)^2 + (219)^2 + \ldots + (224)^2] / 4\} - CF \\
&= 561.33 \\
&= (\Sigma x^2)\text{trts}
\end{aligned}
$$

$$
\begin{aligned}
SSB &= \text{Block sum of squares} \\
&= \{[(318)^2 + (302)^2 + \ldots + (324)^2] / 6\} - CF \\
&= 372.50 \\
&= (\Sigma x^2)\text{blocks}
\end{aligned}
$$

$$
\begin{aligned}
SStot &= \text{Total sum of squares} \\
&= [(60)^2 + (47)^2 + \ldots + (62)^2 + (61)^2] - CF \\
&= 1999.83 \\
&= (\Sigma x^2)\text{total}
\end{aligned}
$$

$$
\begin{aligned}
SSE &= \text{Error sum of squares} \\
&= SStot - (SST + SSB) \\
&= 1999.83 - (561.33 + 372.50) \\
&= 1066 \\
&= (\Sigma x^2)\text{error}
\end{aligned}
$$

Step 3. Perform the ANOVA for yield (Y) to get the Σy^2 values:

$$
\begin{aligned}
CF &= (95.67)^2 / 24 \\
&= 381.36
\end{aligned}
$$

$$
\begin{aligned}
SST &= \{[(19.43)^2 + \ldots + (16.09)^2] / 4\} - CF \\
&= 8.47 \\
&= (\Sigma y^2)\text{treatments}
\end{aligned}
$$

$$
\begin{aligned}
SSB &= \{[(18.34)^2 + \ldots + (29.51)^2] / 6\} - CF \\
&= 10.44 \\
&= (\Sigma y^2)\text{blocks}
\end{aligned}
$$

$$SStot = [(3.42)^2 + (2.87)^2 + \ldots + (5.38)^2 + (5.20)^2] - CF$$
$$= 29.17$$
$$= (\Sigma y^2)\text{total}$$

$$SSE = 29.17 - (8.47 + 10.44)$$
$$= 10.26$$
$$= (\Sigma y^2)\text{error}$$

Step 4. The third ANOVA is for the cross-product XY of each corresponding cell of Step 1, to get the Σxy values:

$$CF = (1310 \times 95.67) / 24$$
$$= 5221.99$$

$$(\Sigma xy)T = \{[(241 \times 19.43) + \ldots + (224 \times 16.09)] / 4\} - CF$$
$$= 34.06$$

$$(\Sigma xy)B = \{[(318 \times 18.34) + \ldots + (324 \times 29.51)] / 6\} - CF$$
$$= 8.96$$

$$(\Sigma xy)tot = \{[(60 \times 3.42) + (47 \times 2.87) + \ldots + (61 \times 5.20)]\} - CF$$
$$= 110.80$$

$$(\Sigma xy)error = 110.80 - (34.06 + 8.96)$$
$$= 67.78$$

Step 5. Perform the analysis of covariance.

Source of variation	d.f.	Σx^2	Σxy	Σy^2	Deviations from regression d.f.	SS	MS	F_c
Total	23	1999.83	110.80	29.17				
Blocks	3	372.50	8.96	10.44				
Cultivars	5	561.33	34.06	8.47				
Error	15	1066.00	67.78	10.26	14	5.95	0.43	
Cultivars + error	20	1627.33	101.84	18.73	19	12.36		
Adjusted means					5	6.41	1.28	2.98

Note: SS = sum of squares
MS = mean squares

The values of the SS of the deviations from regression for error and cultivars + error are calculated as follows:

$$SSdy.x(error) = \Sigma y^2 error - \{[(\Sigma xy)error]^2 / \Sigma x^2 \ error\}$$
$$= 10.26 - [(67.78)^2 / 1066]$$
$$= 5.95$$

$$SSdy.x(cultivars+error) = 18.73 - [(101.84)^2 / 1627.33]$$
$$= 12.36$$

$$d.f. = (t-1) + [(t-1)(b-1)-1]$$

$$SSdy.x(adjusted\ means) = 12.36 - 5.95$$
$$= 6.41$$

$$d.f. = (t-1)$$

The MS values for these sources are estimated by dividing each SS by its corresponding d.f.:

$$MSE = SSE / d.f.$$
$$= 5.95 / 14$$
$$= 0.43 \text{ and}$$

$$MS\ adj.means = SS\ adj.\ means / d.f.$$
$$= 6.41 / 5$$
$$= 1.28$$

The F test for adjusted means is performed by dividing MS adj. means / MS error :

$$F_C = 1.28 / 0.43$$
$$= 2.98$$

$$d.f. = 5.14$$

<u>Step 6.</u> Since there is significance at the 5% level when testing the adjusted means, the covariance is necessary and the mean values for cultivars should be adjusted. The error regression coefficient, byx, is calculated as follows:

$$byx = \Sigma xy \text{ error} / \Sigma x^2 \text{ error}$$
$$= 67.78 / 1066$$
$$= 0.0636$$

Step 7. The adjusted mean treatment values are calculated as follows:

$$\hat{\bar{Y}}_i = Y_i - byx(\bar{X}_i - \bar{X})$$

As an example, for cultivar 1, the adjusted mean treatment value is:

$$\hat{\bar{Y}}_1 = \bar{Y}_1 - byx(\bar{X}_1 - \bar{X})$$
$$= (19.43 / 4) - 0.0636[(241 / 4) - (1310 / 24)]$$
$$= 4.8587 - 0.0636\ (60.25 - 54.58)$$
$$= 4.4981$$

and for cultivar 2,

$$\hat{\bar{Y}}_2 = \bar{Y}_2 - byx(\bar{X}_2 - \bar{X})$$
$$= 14.78 / 4 - 0.0636[(219 / 4) - (1310 / 24)]$$
$$= 3.695 - 0.0636(54.75 - 54.58)$$
$$= 3.684$$

Step 8. Comparison of adjusted treatment (cultivar) means is made individually for each pair of adjusted means since the value of S_d is different for each comparison. Comparison of cultivars 1 and 2:

$$S_{d(1,2)}{}^2 = (MSE)\{(2 / b) + [(\bar{x}_1 - \bar{x}_2)^2 / \Sigma x^2 \text{error}]\}$$
$$= 0.43\ \{(2 / 4) + [(60.25 - 54.75)^2 / 1066]\}$$
$$= 0.2272$$

$$S_{d(1,2)} = (0.2272)^{1/2}$$
$$= 0.4767$$

where b = number of blocks, and the calculated t value comes from:

$$t_c(1,2) = (\hat{\bar{Y}}_1 - \hat{\bar{Y}}_2) / S_{d(1,2)}$$
$$= (4.4896 - 3.6844) / 0.4767$$
$$= 1.6891$$

and since $t_{14df,.05} = 2.145$, then $t_c = 1.6891$ N.S.,

so there is no significance at the .05 level, and cultivars 1 and 2 are not significantly different.

Now, doing the comparison for cultivars 1 and 4:

$$S_{d(1,4)}^2 = 0.43\ [(2 / 4) + (60.25 - 50.5)^2 / 1066]$$
$$= 0.2533$$

$$S_{d(1,4)} = (0.2533)^{1/2}$$
$$= 0.5033$$

$$t_{c(1,4)} = (4.4896 - 3.167) / 0.5073$$
$$= 2.628$$

and since $t_{14df,.05} = 2.145$, then $t_c = 2.628$, and cultivar 1 is significantly different from cultivar 4. The same procedure is followed for all possible comparisons.

III
Exploratory Trials

Exploratory on-farm research is normally conducted at the same time as initial characterization, and as a complement to that process. It can also, however, be used throughout the technology development process. Exploratory trials are useful in at least two types of situations: when research is initiated in a new region, or when no previous information to estimate response to potential new alternatives is available. In new areas, more efficiency can be obtained if diagnostic activities are complemented by exploratory trials. There are two advantages: 1) periodic interaction with farmers provides additional information to complement the diagnosis, and 2) these trials produce valuable information for the design of site-specific and regional trials. In later stages of on-farm research, exploratory trials help redefine or produce new research guidelines, as it is common for good research to generate new questions.

Normally, exploratory trials provide qualitative results that later can be quantified by other types of experiments. Exploratory trials commonly include several (usually three or four and occasionally up to seven or eight) factors, using at least two levels for each factor, with few replications. When available agronomic information is scarce, the number of variables and treatments may be high, and the design may be relatively

complicated. One way to keep down the number of variables in these trials, and keep them small in size, is to design two or more different experiments. By using only three or four variables in each trial and choosing groups of variables that interact frequently (e.g., fertilizer and cultivars, weed control and plant density) the design is further simplified. For the evaluation of potential alternatives, such as introducing a new crop in the region, the number of treatments can be reduced and the design will be simplified. These trials are mostly researcher-managed, though the farmers' previous experience makes their input and opinions in the design of treatments essential. A discussion of the types of exploratory trial designs follows, along with examples for each case.

SUPERIMPOSED TRIALS

A relatively simple, convenient, and efficient means of exploring the effect of different factors in a new area is a superimposed trial. In this type of trial,

TABLE III-1. Example of a superimposed six-treatment N-P-K trial in rice.

Treatment	Grain yield						
	Farm Number						
N P K	1	2	3	4	5	6	$\bar{x}$
	(metric ton x 100)						
50- 0- 0	336	434	451	411	402	375	401.5
90- 0- 0	439	416	506	459	482	431	455.5
70- 0- 0	443	398	457	370	454	350	412.0
70-30- 0	412	419	412	398	499	386	421.0
70-30-30	416	368	482	370	397	402	405.8
70- 0-30	417	377	493	364	490	387	421.3
$\bar{x}$	410.5	402.0	466.8	395.3	454.0	388.5	419.5

Source: Zandstra et al. (1981), p.107.

treatments are placed on fields which are being managed by the farmers themselves. Treatments are marked by stakes

or other means, and individual treatments are installed either by the researcher or the farmer. Together, the researcher and the farmer harvest the crop when it is mature. The design of a superimposed trial should be simple. Replications should be used at each location, although data from designs without replications at each site can be combined for regional analysis and interpretation.

An example of a simple superimposed trial from IRRI is shown in Tables III-1 and III-2. Previous information indicated that rice responded to at least 50 kg/ha of nitrogen, but response to potassium and phosphorus was uncertain. A simple six-treatment superimposed trial was established on a number of farms and information was obtained from six of them (Table III-1). The design at each location was without replication. The six treatments included three levels of nitrogen (50, 70, and 90 kg/ha). At 70 kg/ha of N the treatments explored the application of 30 kg/ha phosphorus and potassium individually and together. Analysis of variance (Table III-2) indicated a significant effect for nitrogen but none for the other elements. The conclusion was that more nitrogen would have a positive effect on yield, and if the cost were less than the value of the additional crop, more nitrogen could be recommended. Furthermore, it would indicate that additional work should be done with nitrogen, but the other major elements (P and K) need not be studied further in this context.

TABLE III-2. Analysis of variance of a superimposed N-P-K rice trial.

Source	d.f.	SS	MS	F_c
Total	35	69,071		
Farms	5	32,178		
Treatments	5	11,212	2,242	
N	2	9,837	4,918	4.79
Remainder	3	1,375	458	0.45
Error	25	25,681	1,027	

CV = 7.6%

Source: Zandstra et al. (1981), p. 107

THE 2^n-FACTORIAL TRIAL

A useful arrangement of treatments that produces exploratory information on several factors and their interaction is a 2^n-factorial trial. This is an arrangement of n factors taken at two levels each. An example of a 2^3-factorial experiment will be used to illustrate the detailed methodology of analysis. The nature of the trial was to explore three factors: plant density, nitrogen, and variety in a 2^3 factorial. The levels of the variables in this trial are given for plant density (P), nitrogen (N), and variety (V) as follows:

P_0 = 25,000 plants/ha
P_1 = 50,000 plants/ha

N_0 = None applied
N_1 = 100 kg N/ha

V_0 = Local variety
V_1 = Tuxpeño

Table III-3 shows the field design by blocks and the grain yield in kg/plot. To estimate the factorial effects and perform the analysis of variance, the following procedure is applied:

Step 1. Using treatment yield totals from Table III-3, a two-way table can be constructed in which the treatments are placed on the horizontal axis and the factorial effects are listed vertically as shown in Table III-4. In each of the factorial effects, half the treatment combinations receive a plus (+) sign and half receive a minus (-) sign. The row corresponding to M (mean of the entire experiment) has only plus signs. When the higher level of a factor (P_1, N_1, V_1) in the factorial effect is present in the treatment combination, it receives a plus (+) sign; it receives a minus (-) sign if not present (P_0, N_0, V_0). When two or more factors at higher levels are present in the treatment combination, the sign is found by using the algebraic rule for signs.

Example: pn (+) = p(-) x n (-)

pnv (-) = p (-) x nv (+)

TABLE III-3. Example of a 2^3 trial on maize with plant density (P), nitrogen (N), and variety (V).

Treatments	Code	Block I	Block II	Treatment totals	Treatment means
		kg/plot			
$P_o N_o V_o$	(1)*	4.3	3.9	8.2	4.1
$P_1 N_o V_o$	p	4.5	5.9	10.4	5.2
$P_o N_1 V_o$	n	4.5	5.4	9.9	4.95
$P_o N_o V_1$	v	5.7	6.6	12.3	6.15
$P_1 N_1 V_o$	pn	6.4	6.7	13.1	6.55
$P_1 N_o V_1$	pv	6.9	7.1	14.0	7.0
$P_o N_1 V_1$	nv	6.4	7.0	13.4	6.7
$P_1 N_1 V_1$	pnv	8.6	8.8	17.4	8.7
Totals		47.3	51.4	98.7	

Source: Adapted from CIMMYT data.
*(1) is the local or traditional treatment or the lowest level combination.

Table III-4 has the following characteristics: 1) every row has an equal number of plus and minus signs, except for row M; and 2) the sum of products of signs in any given pair of rows is zero (Montgomery, 1976). For another reference on the signs for these contrast coefficients see Cochran and Cox (1957, p 157).

Step 2. The total factorial effects are calculated by the algebraic addition of the yields of the same treatment using the corresponding plus (+) or minus (-) sign of each component, and the result is written in the space for the

total of each treatment (Table III-4).

Example:

Total factorial effect for (p):

$p = (P_1\ N_o\ V_o)$
$= -8.2 + 10.4 - 9.9 - 12.3 + \ldots + 17.4$
$= 11.1$

Total factorial effect for (pn) is:

$pn = 8.2 - 10.4 - 9.9 + \ldots + 17.4$
$= 3.3$

Step 3. In the analysis of variance (Table III-5), since there are 16 observations (eight treatments and two replications), the square of the total of each factorial effect is divided by 16 to obtain the sum of squares (SS) for each treatment.

Example: SS for (p) $= (11.1)^2 / 16$
$= 7.7$

The total sum of squares (SStot), the sum of squares for blocks (SSB), and the sum of squares for the error (SSE) are calculated in the usual form.

TABLE III-4. Calculation of the total factorial effect in the 2^3 factorial with two replications.

	Treatment combination: code and total yield (kg)								
Factorial effect	(1) 8.2	p 10.4	n 9.90	v 12.3	pn 13.1	pv 14.0	nv 13.4	pnv 17.4	total 98.7
M	+	+	+	+	+	+	+	+	
P	-	+	-	-	+	+	-	+	11.1
N	-	-	+	-	+	-	+	+	8.9
V	-	-	-	+	-	+	+	+	15.5
PN	+	-	-	+	+	-	-	+	3.3
PV	+	-	+	-	-	+	-	+	0.3
NV	+	+	-	-	-	-	+	+	0.1
PNV	-	+	+	+	-	-	-	+	1.3

$$\text{Correction factor (CF)} = (\text{Grand total})^2 / n$$
$$= (98.7)^2 / 16$$
$$= 608.856$$

where n = number of observations

Total sum of squares (SStot):

$$SStot = \Sigma(\text{each observation})^2 - CF$$
$$= [(4.3)^2 + (4.5)^2 + \ldots + (7.0)^2 + (8.8)^2] - CF$$
$$= 639.45 - 608.856$$
$$= 30.594$$

Sum of squares for blocks:

$$SSB = [\Sigma(\text{each block})^2 / t] - CF$$
$$= \{[(47.3)^2 + (51.4)^2] / 8\} - CF$$
$$= 609.906 - 608.856$$
$$= 1.05$$

where t = number of treatments

The sum of squares for treatments (SST) is calculated in a different manner because the factorial effects have already been calculated (Table III-4).

Sum of squares for treatments (7 treatment combinations):

$$SST = \Sigma(\text{each factorial effect})^2 / n$$
$$= [(11.1)^2 + (8.9)^2 + (15.5)^2 + \ldots + (0.1)^2 + (1.3)^2] / 16$$
$$= 28.46$$

and $SSE = SStot - (SST + SSB)$
$$= 30.594 - (28.46 + 1.05)$$
$$= 1.084$$

Step 4. In the analysis of variance for a 2^n factorial, one degree of freedom is always assigned to each factor or interaction of factors because the factor effect is calculated by comparing two levels of the factor or interaction and one degree of freedom is lost in the estimation.

Perform the analysis of variance (ANOVA):

TABLE III-5. ANOVA for the 2^3 factorial

Source of variation	Degrees of freedom	Sum of Squares	Mean Squares	F_c
Blocks	(b-1)=1	1.05	1.05	6.77 *
Treatments	(t-1)=7	28.46	4.066	26.23 **
Factor P	1	7.77	7.77	50.129 **
N	1	4.95	4.95	31.935 **
V	1	15.016	15.016	96.88 **
PN	1	.681	.681	4.39 NS
PV	1	.0056	.0056	.036 NS
NV	1	.000625	.000625	.004 NS
PNV	1	.1056	.1056	.68 NS
Error	7	1.084	.155	
Total	15	30.594		

CV = 6.38%

* Significant at 5% level

** Significant at 1% level

Interpretation of Results

By examining the treatment mean yields in Table III-3, and the analysis of variance, Table III-5, it can be determined that each of the three factors individually (plant density, nitrogen, and variety) had a highly significant effect on yield. Out of the three factors, the new variety (V) produced the greatest increase (2.05 kg/plot), and plant density (P) was second with a 1.1 kg/plot increase. Although nitrogen (N) had a significant effect, increasing the yield by 21%, an economic analysis should be conducted to determine whether its application is an economically good choice. (Does the yield increase cover the cost of buying and applying the nitrogen

fertilizer?) On the other hand, variety produces a significant yield increase and its effect on cost is minimal, making it a good alternative to introduce in the production system in which the factors were evaluated. If the plant density change does not require too much additional labor at planting time, this may also be a good alternative to introduce.

Comments on Reducing the Size of the Trial

Generally, many degrees of freedom in this type of design (2^n) are associated with higher-order interactions which are difficult to interpret. If the higher-order interactions (third order and higher) are not considered, the size of the trial would be substantially reduced, keeping some of the advantages of the basic factorial arrangements. In this case, it is advantageous to use the fractional factorial (Cochran and Cox, 1957). Example: with $n = 8$, main effects and first-order interactions can be estimated with only $(1/8) \times 2^8 = 32$ treatments. Equally, main effects and second-order interactions can be estimated with only $(1/4) \times 2^8 = 64$ treatments.

Another way to reduce the number of treatments in a 2^n factorial is to select factors based on their importance. Factors and combinations which are considered of little interest from the biological and economic viewpoint, or those which do not interact, can be eliminated. For instance, if selecting two factors, A and B, the following treatments can be established, with two levels each: A1 B1, A2 B1, A1 B2, and A2 B2. If there is no interaction between the two factors, determined from previous experimental information, the main effect of A corresponds to the average difference between A2 B1 and A1 B1, and the effect of B to the average difference between A1 B2 and A1 B1.

THE "PLUS" TRIAL

Exploratory information on new variables as they relate to existing practices can be obtained by testing, one at a time, a series of alternatives that include the new variables. The following example compares a traditional maize practice with three alternatives. It

consists of four treatments with two replications.

Treatment	Description
Traditional (T)	25,000 plants/ha, no N applied, local variety.
T + density	50,000 plants/ha, no N applied, local variety.
T + nitrogen	25,000 plants/ha, 100 kg/ha of N, local variety
T + variety	25,000 plants/ha, no N applied, Tuxpeño variety

Results of this trial are shown in Table III-6.

TABLE III-6. Maize yield data for a four-treatment "plus" trial.

Treatment	Block I	Block II	Treatment totals	$\bar{x}$
	kg/plot			
Traditional (T)	4.3	3.9	8.2	4.1
T + density	4.5	5.9	10.4	5.2
T + nitrogen	4.5	5.4	9.9	4.95
T + variety	5.7	6.6	12.3	6.15
Totals	19.0	21.8	40.8	5.1

Source: Adapted from CIMMYT data.

After calculating the sums of squares the ANOVA table is shown in Table III-7.

TABLE III-7. ANOVA for maize "plus" trial.

Source of variation	d.f.	SS	MS	F_c
Blocks	1	0.98	0.98	3.3 NS
Treatments	3	4.27	1.423	4.8 NS
Error	3	0.89	0.2967	
Total	7	6.14		

C.V. = 10.68 %

$$CF = [(40.8)^2] / 8$$
$$= 208.08$$

$$SSB = \{[(19.0)^2 + (21.8)^2] / 4\} - CF$$
$$= 209.06 - 208.08$$
$$= 0.98$$

$$SST = \{[(8.2)^2 + \ldots + (12.3)^2] / 2\} - CF$$
$$= 212.35 - 208.08$$
$$= 4.27$$

$$SStot = [(4.3)^2 + \ldots + (6.6)^2] - CF$$
$$= 214.22 - CF$$
$$= 6.14$$

$$SSE = SStot - [SSB + SST]$$
$$= 6.14 - (0.98 + 4.27)$$
$$= 0.89$$

The analysis of variance does not show significant differences at the 5% level, although it is very close to being significant at the 10% level. It seems unlikely with such a low C.V. value that no significant difference has been detected when there is a 50% increase in yield when the new variety is used. This is the type of problem encountered in performing ANOVA when the number of degrees of freedom for the error is small, making the F_c values high and difficult to surpass. One solution would have been to increase the number of replications to five to obtain more degrees of freedom for the error.

If a "t" test is performed for the two treatments with the largest difference (e.g., traditional vs. T + new variety), a procedure that is not statistically orthodox, the following would result:

$$t = (\bar{x}_1 - \bar{x}_2) / S_{\bar{x}_1 - \bar{x}_2}$$

$$= (6.15 - 4.1) / 0.4924$$

$$= 4.1629$$

There are 2(r-1) = 2 degrees of freedom, where r = number of replications or blocks. The calculated t values for 2 degrees of freedom are 2.920 for a 10% level of significance and 4.303 for a 5% level. Hence, in this case, the use of the new variety is not significant at the 5% level, but there is a strong indication that it does make a difference in yield under these otherwise traditional practices. Because the seed cost of maize is a relatively small proportion of total production cost, this may still be worth testing further.

The step by step procedure for calculating the above t value is as follows:

$$S_{\bar{x}_1 - \bar{x}_2} = (2S_{\bar{x}}^2 / r)^{1/2}$$

$$= [2(0.2425) / 2]^{1/2}$$

$$= 0.4924$$

and

$$S_{\bar{x}}^2 = (S_1^2 + S_2^2) / r$$

$$= (0.08 + 0.405) / 2$$

$$= 0.2425$$

and

$$S_1^2 = \Sigma x_1^2 / (r-1)$$

$$= 0.08 / 1$$

$$= 0.08$$

$$\Sigma x_1^2 = \Sigma X_1^2 - [(\Sigma X_1)^2 / r]$$

$$= [(4.3)^2 + (3.9)^2] - [(8.2)^2 / 2]$$

$$= [18.49 + 15.21] - (67.24) / 2]$$

$$= 0.08$$

$$S_2^2 = \Sigma x_2^2 / (r - 1)$$
$$= 0.405 / 1$$
$$= 0.405$$

$$\Sigma x_2^2 = \Sigma X_2^2 - [(\Sigma X_2)^2 / r]$$
$$= [(5.7)^2 + (6.6)^2] - [(12.3)^2 / 2]$$
$$= 0.405$$

THE "MINUS" TRIAL

This is the opposite of a "plus" trial. It compares a technological package with alternatives that reduce the package by one variable at a time. An example consisting of three factors follows:

Treatment	Description
Tech. pack. (TP)	Tuxpeño variety; 100 kg/ha of N, 50,000 plants/ha
TP - variety	local variety; 100 kg/ha of N; 50,000 plants /ha
P - nitrogen	Tuxpeño variety; no nitrogen; 50,000 plants/ha
TP - plant density	Tuxpeño variety; 100 kg/ha of N; 25,000 plants/ha

Results of the trial are shown in Table III-8.

TABLE III-8. Maize yield data for a four-treatment "minus" trial.

Treatment	Block I	Block II	Treatment totals	$\bar{x}$
	kg/plot			
Tech. pack. (TP)	8.6	8.8	17.4	8.70
TP - variety	6.4	6.7	13.1	6.55
TP - nitrogen	6.9	7.1	14.0	7.00
TP - plant density	6.4	7.0	13.4	6.70
Totals	28.3	29.6	57.9	7.24

Source: Adapted from CIMMYT data.

After calculation of the sums of squares for each of the sources of variation, the ANOVA table is as follows:

TABLE III-9. ANOVA of a maize "minus" trial.

Source of variation	d.f.	SS	MS	F_c
Blocks	1	0.2113	0.2113	11.79 NS
Treatments	3	5.91375	1.9713	110.02 **
Error	3	0.0538	0.0179	
Total	7	6.17875		

C.V. = 1.85%

** Significant at 1% level.

$$CF = (GT)^2 / n$$
$$= (57.9)^2 / 8$$
$$= 419.05125$$

$$SSB = \{[(28.3)^2 + (29.6)^2] / 4\} - CF$$
$$= 0.21125$$

$$SST = \{[(17.4)^2 + \ldots + (13.4^2)] / 2\} - CF$$
$$= 5.91375$$

$$SStot = [(8.6)^2 + \ldots + (7.0)^2] - CF$$
$$= 6.17875$$

$$SSE = SStot - (SSB + SST)$$
$$= 6.17875 - (0.21125 + 5.91375)$$
$$= 0.05375$$

The highly significant effect among treatments indicates that a mean separation test should be performed to look for statistical differences. For this specific case Duncan's New Multiple Range test (Little and Hills, 1978; Steel and Torrie, 1980) is used.

First start by computing least significant ranges (Rp) by the following formula:

$$Rp = q_a S_{\bar{x}}$$

where q_a = significant "studentized" ranges taken from a table (Steel and Torrie, 1980).

a = level of significance (.05)
p = number of treatment means involved in the comparison
= 2,3,4

The values are summarized below:

	p		
	2	3	4
$q_{.05}(p,3df)$	4.50	4.50	4.50
and R_p is	0.426	0.426	0.426

NOTE: the values for q_a in the table (see Steel and Torrie, p. 586) are taken for
p = 2, 3, 4 and
d.f. = 3

The values for Rp were calculated by using the formula previously shown, where

$$S_{\bar{x}} = (S^2 / r)^{1/2}$$
$$= (MSE / r)^{1/2}$$
$$= (.0179 / 2)^{1/2},$$
$$= 0.0946$$

Then, a summary for the test is shown by calculating the mean separation. First the treatments are listed in order of decreasing means.

Treatment	kg/plot
$\bar{x}_{tp}$	8.7 a
$\bar{x}_{tp}$ - nitrogen	7.0 b
$\bar{x}_{tp}$ - plant density	6.7 bc
$\bar{x}_{tp}$ - variety	6.55 c

Each Rp value is compared to the observed difference between two means. Here it is important to take into consideration the range of number of means involved. For example, to compare $\bar{x}_{tp}$ with $\bar{x}_{tp-v}$ calculate 8.7 - 6.55 = 2.15. This is compared to Rp = 4.26, which would be in the column p = 4, because $\bar{x}_{tp-variety}$ is the fourth mean in the range starting at $\bar{x}_{tp}$. The mean separation is shown above.

The results of this test show that the "tech pack" treatment is significantly different from the other three treatments (95% level of probability). This means that the lack of any of the three factors that form the "tech pack" will cause a significant reduction in the yield of maize under the conditions of this trial.

IV
Site-Specific Trials

Experiment station trials are often designed to search for "potential" or maximum effect of a technology. Experimental cultivars, for example, are frequently screened under conditions which do not limit their expression of genetic potential. This potential, however, is measured only for the one location -- the experiment station. To obtain more useful information, two or more farm locations can be used with the same type of experimental design and analysis in order to measure "deviations from potential" independently at different locations. This type of trial is called a "site-specific" trial.

Because they are usually complex, with a relatively large number of treatments and replications, site-specific trials are only conducted in a limited number of locations. Information sought is agronomic and not socioeconomic, so plots are small. Many possible sources of variation, such as soil fertility, are frequently controlled at the same levels found or used on the station. Farmers' participation is minimal in these researcher-managed trials.

BASIC CONSIDERATIONS

Results from site-specific trials should define a limited range of alternatives to be evaluated regionally before a technology can be passed on for farmer-managed trials.

When designing site-specific trials it is important to keep the recommendation domain concept in mind in order to make the resulting data usable for regional interpretation. For example, all site-specific trials in one recommendation domain should have equal treatments, replications, and plot size. This allows researchers to combine data for regional interpretation (see next chapter).

Plot Size

Plot size must be adequate to achieve trial requirements. While requirements vary from trial to trial, the size of experimental units must both fulfill research requirements and be adapted to practical circumstances. What is desirable must also be in balance with what is possible; common sense must guide the team's work.

It is important for the number of replications to be the same with large or small plots. There is a tendency to believe that larger experimental units make "better" trials and therefore fewer replications are needed. This is not true. Larger plots will increase the cost of trials and they will also increase the probability of a larger experimental error due to heterogeneity within the blocks. In general, plot size will be limited by the amount of land available for the trial on the farm and by the amount of labor or inputs of other resources available during the experiment.

Variety Evaluation

Testing improved genetic material is common for research in farmers' fields. The following five considerations are important in variety testing:

1) Control treatments should include the recommended variety for the region as well as one or more local materials used by farmers. Comparison of experimental varieties against these standards helps to make more meaningful recommendations. The on-farm researcher will not be interested in identifying only the highest-yielding cultivar, but will also be interested in other agronomic characteristics of interest to the farmer.

2) The farmers' own agronomic practices should be strongly respected. The main objective of on-farm evaluation of new varieties is to know their real potential under farmers' conditions. Therefore special "experiment station" handling of these trials should be avoided.

3) Experimental varieties selected for testing should include all available alternatives with a theoretical potential of excellence. This means that not only the experimental varieties of the official research sector should be tested along with local materials, but also varieties from private research programs and from national or international centers and seed companies should be considered.

4) Randomized complete blocks is the experimental design most often appropriate for these types of experiments.

5) The experimental unit should be protected from environmental bias coming from growth habits of neighboring varieties. In maize, for example, where varieties may differ widely in plant size, extra rows of the same variety at each side of the experimental unit should be added. Those border rows are not harvested for experimental purposes. A common practice in maize is to plant four rows of each variety but only use the inside two rows to constitute the experimental unit.

Crop Associations

A common practice among small farmers is to grow two or more crops in the same area. Different crop combinations, row spacing, management, and planting

sequences are common. When typical farmer practices are to be included, a superimposed field trial may be appropriate. On the other hand, when alternatives are dramatically different from typical practices, conventional field trials should be defined. A split-plot arrangement can be appropriate when working with more than one variable. For example, when one variable requires different row arrangements, or if there is a large border effect and the experimental unit size is large, it can be assigned to the main plot. The other variables, such as planting distance, varieties, or secondary crop alternatives can then be assigned to the subplots. Precision will be greater for the variables in the subplots because more degrees of freedom are associated with subplot than with main plot error. An economic interpretation of these types of trials is mandatory since the crops involved generally have different market values, making yield relatively less important (see Chapter VII).

Plant Nutrition

Fertilizer trials are commonly conducted as site-specific experiments. Information on soil characteristics, previous management, and soil analysis should be determined before locating the experiment. Generally, at least three levels of each factor should be considered in order to estimate a response curve. Experimental designs should allow for measurement of interaction effects which are common in fertilizer trials. Factorial designs arranged in randomized complete blocks (RCB) offer a better estimate of interactions among factors than split-plot arrangements. The reason is that, in analysis of variance for the RCB design, the error mean square (MSE) is estimated with more degrees of freedom. The split-plot design has the same number of degrees of freedom for interactions as the RCB, but the residual degrees of freedom have to be distributed between the main plot error and the subplot error.

Special care must be used in field design to avoid fertilizer runoff effects from adjoining plots. Border rows or ample distance should be considered between experimental units. When the local practice is <u>not</u> to use fertilizer, the check plot should reflect that practice. When farmers' practices include <u>some</u> fertilizer use, the

check plot should not be an absolute check, but should reflect the common practice.

Plant Protection

Evaluation of pest (insect, weed, and disease) problems is more difficult than the other agronomic trials discussed. The main reason is that causal agents vary in intensity and mode of action, not only from year to year, but also within a small area. Therefore, pest protection trials require large experimental units with many replications, repeated for various cycles. A factorial arrangement in randomized complete blocks or split-plot designs is convenient. Superimposing the trial on a farmer's field is also a logical option.

The probability distribution of pest damage does not commonly assume a normal pattern. Sample data need to be transformed in order to approximate a normal distribution, which is a theoretical requisite for common statistical methodology. The most frequent transformations for these kinds of data are logarithmic [log X or log (X + 1) when zero values are present]; square root [of X, (X + 1) or (X + 1/2)]; and the angular transformation ARCSINE $(\%)^{1/2}$ when data are given in percentage values between zero and 20 or 80 and 100.

ANALYSIS OF RESULTS

The randomized complete block design is perhaps the most common for site-specific trials. When repeated in other locations, the results can be grouped for a combined analysis, allowing a more meaningful interpretation of site-specific trials.

The example selected to illustrate the statistical methodology is a randomized complete block design that comes from the Chimaltenango area in Guatemala. The experiment is a maize variety trial that was established following a recommendation from the previous year to compare the farmers' own varieties of the region with varieties bred and selected at the local experiment station. It had been concluded in exploratory trials that the station varieties interacted with environment in the region, in many cases yielding less than the local

farmers' own seed selections. Therefore, good varieties selected by the farmers and identified by the FSR/E team were compared with the station varieties.

The farmers' varieties, most of them identified by individual farmers' names, and four station varieties are presented in Table IV-1.

TABLE IV-1. Description of 15 flint cultivars included in site-specific trials of maize in Chimaltenango, Guatemala.

Varieties	Color of endosperm	Code
Los Pitos	White	1
Garcia	Yellow	2
Cojobal	White	3
Ajquejay	Yellow	4
Lopez	Yellow	5
Unec	Yellow	6
V-304 1/	White	7
Argueta	Yellow	8
Marrin	Yellow	9
Santizo	Yellow	10
Tsut	White	11
Ordoñez	Yellow	12
Don Marshall 1/	Yellow	13
Sintetico Chanin 1/	Yellow	14
Chanin-4 1/	Yellow	15

1/ Experiment Station Varieties
Source: ICTA, Guatemala

In Table IV-2, the standardized field values are presented for treatments 1, 2, and 15 to illustrate the methodology for the analysis of variance.

TABLE IV-2. Standardized field data (partial) for yield of maize varieties in four replications in Chimaltenango, Guatemala.

Cultivar	Blocks				Total	$\bar{X}$
	I	II	III	IV		
	metric ton/ha					
1	4.20	5.04	4.90	5.10	19.24	4.81
2	5.20	5.54	5.30	5.80	21.84	5.46
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
15	2.98			3.00	10.36	2.59
Totals	56.97			72.18	265.20	4.42

The sums of squares for each source of variation and the correction factor are calculated as follows:

$$CF = (\text{Grand total})^2 / n$$
$$= (265.20)^2 / 60$$
$$= 1172.184$$

$$SStot = [(4.2)^2 + (5.04)^2 + \ldots + (2.98)^2 + \ldots + (3.00)^2] - CF$$
$$= 1192.564 - 1172.184$$
$$= 20.380$$

$$SSB = \{[(56.97)^2 + \ldots + (72.18)^2] / 15\} - CF$$
$$= 1173.293 - 1172.184$$
$$= 1.109$$

$$SST = \{[(19.24)^2 + (21.84)^2 + \ldots + (10.36)^2] / 4\} - CF$$

$$= 1180.434 - 1172.184$$
$$= 8.250$$

$$SSE = (SStot) - (SSB + SST)$$
$$= 20.380 - (1.109 + 8.250)$$
$$= 11.021$$

Then the degrees of freedom (d.f.) are estimated for each source of variation, considering four replications ($r = 4$) and 15 treatments ($t = 15$). With this information, calculate mean squares by dividing the sum of squares for each source of variation by its corresponding degrees of freedom:

$$MSB = SSB / (r-1) = 1.109 / 3 = 0.3697$$

$$MST = SST / (t-1) = 8.250 / 14 = 0.5893$$

$$MSE = SSE / [(r-1)(t-1)] = 11.021 / 42 = 0.2624$$

Now calculate F values (F_c) by dividing the MS of the sources of variation by the MS of the error.

$$F_c \text{ (blocks)} = MSB / MSE = 0.3697 / 0.2624 = 1.4089$$

$$F_c \text{ treatments} = MST / MSE = 0.5893 / 0.2624 = 2.2458$$

The coefficient of variation of the experiment is calculated from the square root of the MSE, and the general mean of all observations:

$$CV = [(MSE)^{1/2} / \bar{x}](100) = [(0.2624)^{1/2} / 4.42](100) = 11.59\ \%$$

TABLE IV-3. Analysis of variance for yield of 15 maize varieties with four replications, Chimaltenango, Guatemala.

Variation	d.f.	SS	MS	F_c	$F_{.05}$
Blocks	3	1.109	0.3697	1.4089 N.S.	2.83
Treatments	14	8.250	0.5893	2.2458 *	1.94
Error	42	11.021	0.2624		
Total	59	20.3800			

Source: ICTA, Guatemala
CV = 11.59% N.S.: Not significant
*: Significant at 5% level

The calculated F values (F_c) are compared with the F values from a table to determine levels of significance. For blocks there are 3 and 42 degrees of freedom, and for treatments 14 and 42 degrees of freedom. The statistical significance obtained for treatments indicates that yield of at least one of the varieties differs from the rest. In order to further define which varieties are statistically different, a Tukey multiple range test for comparison of means is performed (Steel and Torrie, 1980). This method consists of computing a difference (D), which would be significant at the 5% level, and comparing it with the differences between each pair of treatment means in the experiment. If a difference between two means is equal to or greater than the value for D, then the two means are significantly different.

The calculations are as follows:

$D = Q\ S\bar{x}$ where:

Q = Value which is taken from a statistical table (see Steel and Torrie, 1980, pp. 588-589), and which is a function of number of treatments and degrees of freedom of the error.

and $S\bar{x} = (MSE / r)^{1/2}$
$= (0.2624 / 4)^{1/2}$
$= 0.256$
where r = number of replications

Then $D = 5.11 \times 0.256$
$= 1.31$

The 15 varieties are then grouped in descending order of yields, as shown in Table IV-4. Taking each treatment mean, all possible comparisons are made starting from the top and working down. From the largest mean value (5.52) the value of D (1.31) is subtracted to find all mean yields which are not different from the largest.

TABLE IV-4. Mean yields of 15 varieties of maize evaluated in four replications, Chimaltenango, Guatemala.

CODE	Variety	ton/ha				
3	Cojobal	5.52	a			
2	Garcia	5.46	a			
9	Marroquin	5.01	a	b		
5	Lopez	4.93	a	b		
4	Ajquejay	4.83	a	b		
1	Los Pitos	4.81	a	b		
6	UNEC	4.80	a	b		
7	V-304	4.72	a	b		
10	Santizo	4.57	a	b		
8	Argueta	4.55	a	b		
11	Tsut	4.54	a	b		
12	Ordoñez	4.00		b	c	
13	Don Marshall	3.14			c	d
14	Sintetico Chanin	2.92			c	d
15	Chanin-4	2.59				d

Source: ICTA, Guatemala
CV = 11.59; $S\bar{x} = 0.256$; $\bar{\bar{x}} = 4.42$

The range is indicated by the letter "a" at the right. The range for the next highest yield is from 5.46 to (5.46 - 1.31) = 4.15, which includes the same cultivars.

The range for the third highest yield is from 5.01 to (5.01 - 1.31) = 3.70, which includes all but the lowest three indicated by the letter "b". The letters at the right of the treatment means in Table IV-4 indicate the four groups within which there is no significant difference.

Table IV-5 presents the agronomic characteristics of the varieties under evaluation.

TABLE IV-5. Agronomic characteristics of 15 varieties evaluated in a site-specific trial in Chimaltenango, Guatemala.

Variety	Days to Flower	Plant Height (m)	Ear Height (m)
1. Los Pitos	129	2.85	1.60
2. Garcia	127	3.25	2.40
3. Cojobal	127	3.35	2.20
4. Ajquejay	126	3.10	2.15
5. Lopez	128	3.45	2.30
6. Unec	127	3.15	2.90
7. V-304	117	2.25	1.15
8. Argueta	124	2.55	1.25
9. Marroquin	120	3.40	2.25
10. Santizo	119	2.50	1.45
11. Tsut	128	3.55	2.40
12. Ordoñez	129	3.40	2.25
13. Don Marshall	107	2.00	1.00
14. Sintetico Chanin	102	2.20	1.15
15. Chanin-4	103	2.70	1.20

Source: ICTA, Guatemala

The analysis shows the highest-yielding group of varieties among which there is no significant difference (varieties 3, 2, 9, 5, 4, 1, 6, 7, 10, 8, and 11), and two (varieties 3 and 2) which are significantly different from the four lower-yielding varieties (12, 13, 14, and 15). The best varieties were the local materials and the lowest-yielding were the experiment station varieties Don Marshall, Sintetico Chanin, and Chanin-4. These last varieties were selected for earliness and lower plant height.

Farmer evaluation of the materials in this trial indicated that shorter plants were not generally desired because corn stalks are used for fences and walls. However, the station varieties, although lower-yielding and more than one meter shorter, proved to be nearly 30 days earlier than most local varieties. It was concluded that these characteristics offer good potential for alternative cropping systems (intercropping, relay cropping, consecutive cropping).

V
Researcher-Managed Regional Trials: Agronomic Evaluation

Regional trials are a set of similar trials conducted in a region previously identified as a recommendation domain. Their main objective is the evaluation of data from on-farm and on-station trials to define the interaction of technology with environmental conditions, both from an agronomic and a socioeconomic viewpoint. Verification of homogeneity within the previously identified recommendation domain may result, or evidence supporting the necessity to partition the recommendation domain can be obtained. Recommendations for treatments (technologies) to be submitted to farmer-managed trials should result from analysis and interpretation of regional trials.

In designing regional trials, the number of locations should be as high as resources permit, with the regional experiment station serving as one site. In a single recommendation domain there should probably be no fewer than five locations. Analyses can be made with fewer locations, but precision will be questionable.

Farmers should participate in the management of the trials with full knowledge of the variables studied and the results expected. Throughout the experiment, farmers should be in close contact with the person or persons responsible for the trials. Farmers' active participation adds resources and reduces the required researcher input at each location; it therefore

facilitates the use of more locations.

If a sufficient number of locations have been used for site-specific trials, and if they have designs and treatments in common, they can be analyzed as regional trials -- a cost-effective utilization of information. Usually though, a trial will be designed especially for regional analysis with fewer treatments than typical site-specific trials but with a design common for all sites. Variables included in regional trials, then, can be the same as those that were included in site-specific trials, a subset of them, or others based upon different criteria.

The methodology of combining an analysis of variance of data from all locations permits a measurement of the interaction of technology with environment. It also allows for a statistical interpretation of the relative stability of each technology by a partitioning of the total degrees of freedom due to treatments, and for utilizing regression techniques involving environmental indexes (see Chapter VII for a description of modified stability analysis).

Farmers' participation in these trials contributes to the researchers' focus on the farmers' reality, allowing adjustments in experimental design and generating conclusions that would not be possible from a strict numerical interpretation of resulting trial data. For example, farmers could readily reject the color or shape of an experimental bean cultivar in a variety trial, or, in the case of maize, point out the inadequacy of husk coverage, or the impossibility of a suggested thinning practice because of local religious beliefs.

DESIGN AND METHODOLOGY

Technologies to be evaluated in regional trials usually are selected on the basis of results from exploratory and site-specific trials conducted during the previous year or years. Perhaps such trials were concerned with individual components, such as variety, fertilizer, or insecticide. In regional trials these may be combined into a more comprehensive system. From all previous trials in a region, a consensus is formed by a multidisciplinary team as to what factors need to be researched on a broader basis.

The general approach is to conduct a set of trials having standard experimental and treatment design, plot size, and number of replications throughout the region. A brief description of some of the more important choices to be made in designing these experiments follows.

Experimental Designs

A randomized complete block design is preferred, because of its simplicity and precision. Split-plot or Latin square designs are also possibilities, but they may be unnecessarily complex. A separate randomization should be carried out for each block of the trial at each site. In other words, a standard randomization should not be used for all sites.

Number of Replications

To provide an estimate of the experimental error from each site, replicates within sites are necessary. Three or four randomized blocks are recommended, although in extremely limited land situations two blocks per site could be used if compensated for by more sites. For highly variable conditions (such as plant disease control experiments), more than four replications may be required.

Number of Treatments

In order to keep the land area small and to limit the complexity of the trials from a management point of view, the number of treatments should be as small as possible and not exceed 15 to 20. As an example, a complete 3^3 factorial N-P-K rate trial would require 27 treatments, which would be too many. The number of treatments can be reduced by such techniques as confounding to generate incomplete factorials or using other appropriate treatment designs (double square, central composite, etc.).

Control Treatments

Specific treatments to include in a set of trials will depend upon the factors being studied and the combinations of levels needed. Whenever possible, regional trials should include the following treatments as controls:

1) At each site the individual farmer's own technology for the crop;
2) Technology representative or typical for the crop in the recommendation domain; and
3) The currently recommended technology for the recommendation domain.

The first control is to give each farmer a basis for comparison and to provide researchers an estimate of experimental bias. (Is yield within the experimental area at each site greater or less than the farmer's yield?) The second control compares the typical practice in the region with the other treatments. This second control, when used unchanged year afer year, serves also as a benchmark to evaluate the year effect for trials conducted over time. A third control, representing the current recommendation, is included to see how the new technology being studied compares. The second and third controls provide control conditions for all remaining treatments. If the treatments are considered additive to present practices, the second control provides these conditions. If the treatments are considered as additive to recommended practices, the third control provides these conditions.

Kinds of Treatments

Analysis of the data should be anticipated when choosing treatments. Care should be taken to assure that necessary comparisons can be made readily and that differences may be found if indeed they exist. For quantitative variables, the total range and the spacing between levels should be carefully chosen to assure that the treatment range will provide the response desired and that regression may be estimated with adequate precision. Equal spacing of the levels, although recommended and convenient from a statistical point of view, is not always

necessary.

Two multi-location examples are given to show what types of treatments might be chosen in particular situations and how these treatments might relate to one another. The first example is a set of trials at ten locations to evaluate five varieties ($V_1, \ldots V_5$) each at three nitrogen fertilizer rates (N_1, N_2, N_3), as shown in Table V-1. The major portion of these fifteen treatments is a 5 x 3 factorial. In addition to these 15 treatment combinations, three controls are included.

TABLE V-1. Treatment combinations (5 X 3 factorial + 3 controls).

$V_1N_0CP_c$	$V_4N_0CP_c$
$V_1N_1CP_c$	$V_4N_1CP_c$
$V_1N_2CP_c$	$V_4N_2CP_c$
$V_2N_0CP_c$	$V_5N_0CP_c$
$V_2N_1CP_c$	$V_5N_1CP_c$
$V_2N_2CP_c$	$V_5N_2CP_c$
$V_3N_0CP_c$	$V_fN_fCP_f$
$V_3N_1CP_c$	$V_cN_cCP_c$
$V_3N_2CP_c$	$V_rN_rCP_r$

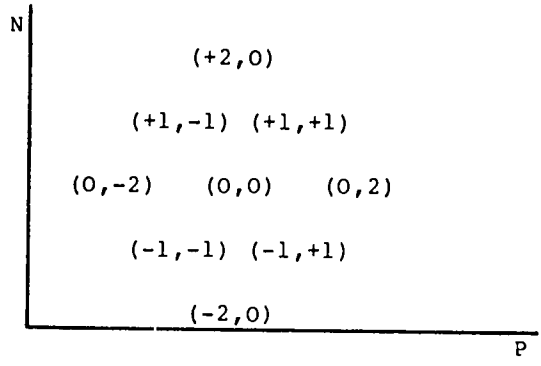
The controls represent:

1) <u>each farmer's</u> variety (V_f), nitrogen fertilizer rate (N_f), and cultural practices (CPf);
2) the variety (V_c), nitrogen rate (N_c), and cultural practices (CP_c) <u>typical</u> of the area, or those commonly used by most farmers; and
3) the variety (V_r), nitrogen rate (N_r), and practices (CP_r) currently being <u>recommended</u> for the region.

In this trial, the variety and fertilizer treatments are considered additive to the common technology of the area (CP_c), so those practices which constitute that technology are used in the major portion of the treatment set.

A second example illustrated in Fig. V-1 represents a fertilizer trial in which the intention is to measure the response to nitrogen (N) and phosphorus (P) over a range of rates, and to estimate a yield response surface within this range. It is assumed that a good variety has already

been adopted by the farmers within the region, so this factor will not need to be included within the trial. The ranges in rates to be used are 0 to 100 kg/ha for N and 0 to 120 kg/ha for P.



The coding scale is as follows:

Actual N:	0	25	50	75	100	Actual P:	0	30	60	90	120
Coded N:	-2	-1	0	+1	+2	Coded P:	-2	-1	0	+1	+2

FIG. V-1. Modified central composite design.

The design selected is a modification of the central composite response surface (described by Cochran and Cox, 1957).

Selection of Sites

Sites should normally be selected to cover a range in the environmental characteristics whose interaction with the technologies to be tested is considered to be of interest. These environmental characteristics could include soil nutrient levels, soil moisture, climatic effects, management, etc. The number of sites required will depend upon the variability in the region and how the results will be used (that is, the analysis or analyses to be employed). For example, if modified stability analysis is to be used where an environmental index is estimated, as suggested in Chapter VII, a minimum of eight to ten sites (perhaps as few as five) is necessary. This may also be a reasonable number for estimating the component of variation due to sites by analysis of variance. One

might also consider the magnitude of differences that one wishes to measure, and then adjust the number of sites so that the error variance component is small enough to detect such a difference.

Complexity

It is desirable that these trials not be highly complex. This is one reason why use of the randomized complete block design is recommended. As mentioned above, treatments may be arranged with either complete or incomplete factorial structure. Treatments in incomplete factorials should be chosen in such a way that undue difficulty in analysis is avoided. Every attempt should be made to standardize the number of blocks, the treatments, and the plot and block dimensions, and to avoid missing values for any of the plots. Trials that differ from the others with respect to these considerations could complicate the analysis. All management operations should be recorded at each site for use in interpreting results.

Replications over Years

Multi-year testing to ascertain stability of results occurs in the farming systems approach as alternative technologies move through the sequence of researcher- and farmer-managed trials. Evaluating technologies over a range of environments also aids in the evaluation of stability. For this reason, it is not normally necessary to repeat the same trial for two or more years, as is usual on experiment stations, which represent only a single site.

COMBINING DATA OVER SITES

One of the goals in regional trials is to provide an estimate of interaction of sites (environments) and treatments. One way this can be accomplished is with a combined analysis of variance over sites. If this interaction is negligible, estimates of the treatment effects over sites, which would be used for interpretation, would be stable. This would imply that the recommendation domain is homogeneous with respect to

the treatments included and need not be partitioned.

Data for the first of the three control treatments (the individual farmer's own technology) would normally be omitted in a combined analysis of variance because it is somewhat different at each location. The second and third kinds of controls would be comparable from site to site and can therefore be included. In general, a format for an analysis of variance which combines data over sites is as follows:

TABLE V-2. Combined analysis of variance procedure.

Source	d.f.	S.S.	M.S.	F
Site (s)	s-1	SSS	$\frac{SSS}{(s-1)} = MSS$	$\frac{MSS}{MS(SxT)}$
Blocks within sites	s(b-1)	SSB(S)	$\frac{SSB(S)}{s(b-1)} = MSB(S)$	$\frac{MSB(S)}{MSE}$
Treat-ments	t-1	SST	$\frac{SST}{(t-1)} = MST$	$\frac{MST}{MS(SxT)}$
Site x treat-ment	(s-1)(t-1)	SS(SxT)	$\frac{SS(SxT)}{(s-1)(t-1)} = MS(SxT)$	$\frac{MS(SxT)}{MSE}$
Error	s(b-1)(t-1)	SSE	$\frac{SSE}{s(b-1)(t-1)} = MSE$	
Total	sbt-1			

Combined Analysis of Variance

The combined analysis of variance and a test for multiple comparison of treatment means is the procedure often chosen to evaluate technologies across a region and/or over time. The new technologies under analysis are very often new varieties or hybrids, but any other

TABLE V-3. Plot yields of 15 maize cultivars evaluated in three sites by ICTA, Chimaltenango, Guatemala

Site	Cultivar	Replications I	II	III	IV	Totals
Alameda	1	2.10	2.27	2.34	2.67	9.38
	2	2.54	2.56	1.94	2.42	9.46
	3	3.16	3.29	2.63	2.36	11.44
	4	2.39	2.56	3.20	2.40	11.55
	5	3.65	3.24	3.25	2.40	12.54
	6	4.65	3.42	2.48	2.42	12.97
	7	3.43	3.44	3.02	3.42	13.31
	8	2.12	3.13	3.64	2.41	11.30
	9	3.39	3.73	3.46	3.64	14.22
	10	2.44	2.16	1.90	1.27	7.77
	11	4.67	3.14	3.15	2.87	13.83
	12	2.61	3.40	3.90	2.84	12.75
	13	1.95	3.44	2.99	3.49	11.87
	14	2.67	3.16	2.58	2.56	10.97
	15	3.33	3.13	2.18	2.59	11.23
	Site totals	45.10	47.07	42.66	39.76	174.59
Parramos	1	4.59	6.26	6.32	4.64	21.81
	2	3.86	6.24	4.10	6.01	20.21
	3	4.75	4.21	4.48	4.65	18.09
	4	6.20	7.00	4.57	5.37	23.14
	5	5.97	7.09	4.29	6.09	23.44
	6	6.43	6.09	5.88	6.70	25.10
	7	6.59	6.21	7.19	5.41	25.40
	8	5.24	6.17	6.33	7.74	25.48
	9	6.40	6.68	5.56	6.00	24.64
	10	3.46	4.08	2.89	3.06	13.49
	11	6.31	6.19	6.52	7.13	26.15
	12	5.03	6.36	4.80	4.10	20.29
	13	6.92	5.85	6.49	6.40	25.56
	14	5.95	6.32	7.30	7.23	26.80
	15	5.35	5.33	4.99	3.64	19.31
	Site totals	83.05	90.08	81.71	84.17	339.01
Itzapa	1	2.56	3.40	3.68	2.22	11.86
	2	3.51	3.71	3.22	3.37	13.81
	3	3.78	3.15	3.27	3.48	13.68
	4	3.62	2.65	3.25	3.39	12.91
	5	3.64	2.39	2.37	2.76	11.16
	6	3.26	3.26	3.29	2.98	12.79
	7	1.92	3.68	3.53	2.40	11.53
	8	2.81	3.20	4.04	3.40	13.45
	9	2.90	2.78	2.99	3.04	11.71
	10	1.65	1.38	2.04	1.94	7.01
	11	2.82	2.03	2.75	2.73	10.33
	12	4.15	3.08	2.92	1.74	11.89
	13	2.53	2.42	3.54	2.29	10.78
	14	3.60	3.32	3.16	3.35	13.43
	15	2.08	3.04	2.26	3.33	10.71
	Site totals	44.83	43.49	46.31	42.42	177.05

management practice or technological innovation generated for the region that shows variation due to soil, management, and/or climatic factors could be analyzed by this procedure. Table V-3 shows a group of three trials illustrating the combined analysis procedure. This group of trials was evaluating the response of yield and morphological aspects of 15 maize cultivars to prevalent conditions in a region.

The procedure for the combined analysis of variance for the cultivars across the three sites was shown in Table V-2. Each of the sums of squares (SS) was calculated as follows:

CF = Correction factor
= [(Grand total of trials)^2] / tbs
= [(tot site 1 + tot site 2 + tot site 3)^2] / (15 x 4 x 3)
= (174.59 + 339.01 + 177.05)^2 / 180
= 2649.986

where t = number of treatments
b = number of blocks or replications
s = number of sites

SSS = Sum of squares sites
= {[(tot of observations in each site)^2] / tb} - CF
= {[(Tot site 1)^2 + (tot site 2)^2 + (tot site 3)^2] / (15x4)} - CF
= {[(174.59)^2+(339.01)^2+(177.05)^2] / 60} - 2649.986
= 295.950

SSB(S) = SS blocks (within sites)
= [(tot observations ea/block, ea/site)^2 / t] - SSS - CF
= {[(tot block 1,site 1)^2 +...+ (tot block 4, site 3)^2] / 15} - SSS - CF
= {[(45.10)^2 + (47.07)^2 +...+ (42.42)2] / 15} - 295.968 - 2649.986
= 5.2908

$$
\begin{aligned}
SST &= \text{SS treatments} \\
&= [(\text{tot of each trt across sites})^2 / bs] - CF \\
&= \{[(\text{trt tot 1 all blocks and all sites})^2 + \ldots \\
&\qquad + (\text{trt tot 15})^2] / (4 \times 3)\} - CF \\
&= [[(9.38 + 21.81 + 11.86^2 + \ldots \\
&\qquad + (11.23 + 19.31 + 10.71)^2 / 12\} - 2649.986 \\
&= 41.3328
\end{aligned}
$$

$$
\begin{aligned}
SS(SxT) &= \text{Sum of squares site x treatment} \\
&= \{[(\text{trt tot ea/site})^2] / b\} - SSS - SST - CF \\
&= [[(\text{tot trt 1, site 1})^2 + \ldots \\
&\qquad + (\text{tot trt 15, site 3})^2] / 4\} \\
&\qquad - SSS - SST - CF \\
&= \{[(9.38)^2 + (9.46)^2 + \ldots + (10.71)^2 / 4\} \\
&\qquad - 295.950 - 41.3328 - 2649.986 \\
&= 28.5818
\end{aligned}
$$

$$
\begin{aligned}
SS\ (tot) &= \text{Total sum of squares} \\
&= \Sigma(\text{each observation})^2 - CF \\
&= [(\text{trt 1, block 1, site 1})^2 + \ldots \\
&\qquad + (\text{trt 15, block 4, site 3})^2] - CF \\
&= [(2.10)^2 + (2.27)^2 + \ldots \\
&\qquad + (3.33)^2] - 2649.986 \\
&= 421.049
\end{aligned}
$$

$$
\begin{aligned}
SSE &= \text{Sum of squares error} \\
&= SS(tot) - [SSS + SSB(S) + SST + SS(SxT)] \\
&= 421.049 - (295.950 + 5.2908 + 41.3328 \\
&\qquad + 28.5818) \\
&= 49.8936
\end{aligned}
$$

To obtain the mean squares, each sum of squares is divided by its corresponding degrees of freedom (Table V-4). The F values for sites and treatments are calculated from MS (SxT) and not MSE. The F value for site-by-treatment interaction is calculated from MSE.

$$
\begin{aligned}
\text{For sites: } F &= MSS / MS(SxT) \\
&= 147.975 / 1.021 \\
&= 144.93^{**}
\end{aligned}
$$

For treatments: F = MST / MS(SxT)
= 2.952 / 1.021
= 2.89**

For site-by-treatment interaction: F = MS(SxT) / MSE
= 1.021 / 0.394
= 2.59

All of these F values are highly significant (P<0.01).

TABLE V-4 Combined analysis of variance for yield of the maize varieties evaluated in three sites, Chimaltenango, Guatemala.

Sources of variation	d.f.	SS	MS	F_c	$F_{.01}$
-Sites	(s-1)=2	295.950	147.975	144.93	7.64
-Blocks (within sites)	s(b-1)=9	5.2908	0.588		
-Trts.	(t-1)=14	41.3328	2.952	2.89	2.80
-Sites x trts.	(s-1)(t-1)=28	28.5818	1.021	2.59	1.87
-Error	s(b-1)(t-1)=126	49.5936	0.394		
-Total	(sbt)-1=179	421.049			

Source: ICTA
CV = 16.4%

The treatment means for each site and overall treatment means are presented in Table V-5. As one studies this table, the reason for the very highly significant effect of sites becomes obvious. The mean yields for the cultivars in Site 2 (Parramos) are, in most cases, nearly double those of the other sites, so this site is responsible for the highly significant difference among sites. The effect of the site-by-treatment interaction is evident when studying individual cultivars. As an example, cultivar 2 has a high yield in Site 3 and a low yield in Site 2. Cultivar 13 responds the opposite way. See Chapter VII for a means of further analyzing this interaction.

TABLE V-5. Site and overall mean yields of 15 maize genotypes evaluated in three sites, Chimaltenango, Guatemala.

Treatment (Cultivar)	Site 1 Alameda	Site 2 Parramos	Site 3 Itzapa	Overall Mean
	metric ton/ha			
1	2.345	5.452	2.965	3.587
2	2.365	5.053	3.451	3.623
3	2.861	4.524	3.420	3.601
4	2.888	5.785	3.226	3.967
5	3.136	5.861	2.789	3.928
6	3.244	6.277	3.199	4.240
7	3.326	6.349	2.882	4.186
8	2.824	6.369	3.363	4.185
9	3.553	6.161	2.927	4.214
10	1.942	3.373	1.755	2.357
11	3.460	6.537	2.583	4.194
12	3.186	5.072	2.972	3.744
13	2.969	6.415	2.693	4.026
14	2.742	6.702	3.357	4.267
15	2.808	4.826	2.677	3.437

Source: ICTA, Guatemala

To determine where the differences exist among cultivars, Tukey's multiple range test for means (see Chapter IV) was performed:

$$
\begin{aligned}
D &= Q\ S\bar{x} \\
&= Q_{t,df}\ (S^2 / bs)^{1/2} \\
&= Q_{15,126} \times (0.3958 / 4x3)^{1/2} \\
&= 4.90 \times (.1816) \\
&= 0.89 \text{ ton/ha}
\end{aligned}
$$

The smallest difference between the mean yields of two cultivars that will make them significantly different at the 5% level of confidence is 0.89 ton/ha.

All treatment means are listed in descending order and the same letter is placed next to all those that are not different when compared:

Cultivar	Overall mean (Metric ton/ha)
14	4.267 a
6	4.240 a
9	4.214 a
11	4.194 a
7	4.186 a
8	4.185 a
13	4.026 a
4	3.967 a
5	3.928 a
12	3.744 a
2	3.623 a
3	3.601 a
1	3.587 a
15	3.437 a
10	2.357 b

For this case, Tukey's multiple range test shows that with the exception of cultivar number 10, there are no significant differences in yield. Cultivar 10 is significantly lower in yield.

Tukey's multiple range test for means is considered by many researchers as too severe since differences have to be large to be statistically significant (5% level). For example, in this case there were differences of up to 20% in yield (.85 ton/ha) among the group where no differences were detected. However, with this test there is little risk of adjudging significant differences that actually do not exist. There are a number of multiple range tests for means and the most appropriate should always be chosen. For a discussion on multiple range tests for means refer to Chew (1977). Had two more sites been included in this trial, a modified stability analysis (Chapter VII) may have been useful for detecting superior cultivars.

Agronomic data serve to complement yield information. In this case, when yield shows no significant differences, the agronomic information is particularly important. Table V-6 shows the mean values on the farm sites for days to flower, ear height, and percentage of rotted ears. Varieties are ordered according to yield.

TABLE V-6. Evaluation of days to flower, ear-height, and rotted ears in 15 maize varieties grown in three sites, Chimaltenango, Guatemala.

Treatment (cultivar)	Days to flower	Ear height (cm)	Rotted ears (%)
14	128	117	6.3
6	127	124	6.3
9	130	132	5.0
11	132	135	6.1
7	129	127	6.1
8	134	135	5.8
13	125	105	4.9
4	129	128	6.7
5	129	130	6.9
12	132	118	5.6
2	125	121	2.5
3	130	123	4.8
1	127	125	5.5
15	131	126	7.1
10	127	119	14.2
$\bar{x}$	129	124	6.3
CV%	2.02	9.91	50.4

Source: ICTA, Guatemala

Variety 10 showed the lowest yield and also the highest percentage of rotted ears (14.2%). Other varieties ranged from 2.5% to 7.1%. Days to flower and ear height show very contrasting values. Varieties 11 and 13 are the earliest (128 and 125 days to flower) and also have the lowest ear height (117 and 105 cm). Concomitantly, varieties 11 and 8 are the latest flowering (132 and 134 days) and the tallest (135 and 135 cm).

These observations contributed to the conclusion that these varieties, in view of their contrasting phenotypes, should be considered for further evaluation in regional trials, as well as in exploratory trials designed to identify novel crop associations and plant distribution.

VI
Researcher-Managed Regional Trials: Socioeconomic Evaluation

Economic evaluation of agronomic research data from on-farm trials can involve little more than gathering and using appropriate information. For example, if a change of fertilizers is involved, the researcher should have the price of the fertilizer currently being used and the price of the fertilizer being substituted. The difference between the rate applied, times the price of each, is the difference in cost of the change. Yield differences are evaluated on the basis of the price of the crop. If yields change sufficiently so that harvest costs are modified, this should be taken into consideration also. Other economic analyses, however, can be more complex and may require or benefit from different analytical or statistical techniques. They may also require different experimental designs from those utilized for agronomic evaluation. Some of these differences include choice of evaluation criteria (the means of measuring results), effect of a technology on other enterprises on the farm, effect on farm labor and its distribution within the family, the potential risk involved, and the demand created for other kinds of inputs. Many of these considerations are too complex to be considered in this volume. However, some warrant discussion, since they are of basic importance to interpreting agronomic trials and making recommendations from them.

Previous Page Blank

CHOICE OF EVALUATION CRITERIA

One of the most important and critical decisions to be made when designing research and evaluating data is the selection of choice criteria for evaluation. The most common evaluation criterion used by agronomists is yield per unit of land area, frequently kg/ha. The use of this criterion implies that land is the most limiting resource on the farm and that therefore productivity of the land is the most important evaluation criterion. This is not always the case. On many small farms, even though there is little land, land is not the most limiting constraint. Nor is the same constraint necessarily the most limiting for different crops. For example, small farmers in Nariño, in the south of Colombia, traditionally plant their scarce potato seed by spacing it widely to maximize the productivity per unit of potato seed. The amount of seed determines the size of the potato field. Hence, land is not the most limiting resource with respect to potato production on these small farms. However, the rest of the land on these farms is planted into grain crops. For grain, land is a limiting resource. For this reason, in the case of potatoes, technological changes which increase the productivity per unit of land area but decrease the productivity per unit of seed will not be attractive to these farmers. On the other hand, the same kind of technology for grain crops could be acceptable. The importance of using the relevant choice criterion in evaluating on-farm trials is obvious in this case.

In large areas of Africa, land is not a limiting resource. Farmers can plant as much land as they are able to manage. However, in these same areas, rainfall is scarce, so weeding the crops becomes a critical factor. These farmers tend to plant the amount of land they can effectively weed, because planting more land is a waste of effort if it cannot be weeded. In this case, labor for weeding becomes the most important evaluation criterion and changes in crop production practices must be evaluated against this factor.

In some areas, such as eastern Guatemala, crops must be planted as soon as possible after the initiation of the rains. Delayed planting reduces yield heavily because of a mid-season dry spell, increased pest problems, or because the crop does not mature before the rains

terminate. In this case, labor available for planting becomes the most important evaluation criterion. For most small farmers, cash is a very limiting resource. For these farmers, the needs of the family and the home compete directly with the needs of the crops and livestock for the limited cash resources available. Non-cash inputs are more important on these farms and only a limited use is made of inputs which require a source of cash. In this case a relevant choice criterion is the comparison of return to cash costs.

Labor Input as an Evaluation Criterion

In all of the above cases, it is necessary for the researcher to have information on the use of the resource in question in order to be able to employ the appropriate evaluation criterion. For example, if labor at weeding time is critical, any changes in technology which influence weeding will create a need for the researcher to monitor weeding labor. Weeding of the individual crop in question is important, but the effect of utilizing more labor on any one particular crop can also have an influence on labor availability for other crops, livestock, or household activities. This should also be taken into consideration. Other important considerations are which members of the household are involved and whether or not labor is hired.

In order to evaluate changes in labor requirements from on-farm or on-station research, it is usually necessary to have larger plots than are required for strictly agronomic evaluation. Plots need not be full field size, however. In on-farm research, records must be kept of the labor utilized by farmers on their own fields for the particular operations in question. Labor use should take into consideration the time lost in going to and from the field, and for resting, drinking water, eating, sharpening implements, or anything else that reduces the amount of time that land can be worked in a day. Records should also be kept of the labor required for the same tasks in the trial plots. Usually, a test plot will be completed without any rest, implement sharpening, or other delays which are normal in field work.

PRACTICAL FIELD ADVICE
Farmers have standard estimates of the amount of time required to do most traditional field operations. The amount of land a person can weed, plant, or clear in a day is established as much by custom as by measurement. Where these measurements are available, they should be used for the traditional treatments because it reflects reality in the area. The problem comes when changes in the amount of labor required to accomplish specific tasks are made. If more time is required to weed a certain area, the people who are doing the work may demand more pay. On the other hand, if less time is required, the farmer may try to increase the area which is completed in a day's time. Evaluating these possible changes must be done by asking the opinions of farmers and laborers or by waiting to see what happens in practice.

Having the information from farms and plots, the researcher can make appropriate adjustments in plot data. By converting plot size to hectares or some other measure of land area commonly utilized by farmers, a factor is available for comparing plot labor usage with real farm data. For example, if a 10 x 10 meter plot requires one half hour for an operation, this is equivalent to 50 hours/ha or 6.25 workdays of 8 hours/day. If the same operation on a farm basis requires 20 workdays per hectare, then 20 / 6.25 = 3.2 is the conversion factor. In other words, multiplying labor operations done on a 10 X 10 plot by 3.2 and adjusting plot size to a normal land unit size results in an estimate of the labor required on a normal land unit basis. For example, if labor required for another 10 X 10 m plot is 20 minutes (1/24 workday), then workdays per hectare for this technology is:

$$(1 / 24) \times 3.2 \times 100 = 13.3$$

If labor for weeding is the limiting resource in a recommendation domain, then production should be divided by weeding time to get an estimate of labor productivity (product per unit of labor). For example, consider the

following situation: under traditional practices, 20 workdays are required to weed a hectare of land that produces 3200 kg of product. A new arrangement in planting density and spacing adds 20% to labor requirements for weeding and results in 500 kg of additional production. Traditional practices result in 160 kg per workday spent in weeding (3200 / 20) but the alternative technology yields 3700 kg for 24 workdays or 154.2 kg per workday. Because labor for weeding is limiting, the relevant evaluation criterion is not yield per hectare (3700 vs. 3200 kg). If farmers are limited to 30 workdays of labor in the weeding period, during those 30 workdays they are able to weed less land in the new system. Because each workday results in 154.2 kg of product with the new technology, rather than 160 kg with the traditional technology, the 30 workdays produce only 4626 kg under the alternative technology, compared with 4800 kg under the traditional technology. In this example, it is evident that the criterion of yield per hectare (3700 kg compared with 3200 kg) results in a false conclusion regarding the value of the technology to the farmers. With a limit of 30 workdays of labor available for weeding, this leaves farmers with a choice of 4800 kg of product with the traditional system or 4626 kg using the new technology.

Measures of productivity of labor (kg/workday) are subject to the same kind of variation as other measures of productivity, such as kg/ha. A common mistake in making an economic analysis is to make only one estimate and assume it is firm. Measures of productivity (evaluation criteria) can be subjected to the same kinds of statistical analysis as are commonly used for biological or agronomic criteria. For example, net income can be affected by variation in yield, price of the product, use of inputs, and price of inputs. When farmers' practices and prices paid or received vary, separate income calculations should be made for each farm, just as separate yields are measured for each farm.

Cash as an Evaluation Criterion

In commercialized and monetized agriculture, cash can effectively substitute for most other inputs. If more seed is needed, it is purchased with cash (or credit, which is another form of cash). If more labor is needed, it is also purchased with cash. However, in many small, limited-resource farm situations, nearly all resources used in the production process come from the farm. Only a very few inputs are purchased. These include inputs which are not available on the farm, such as chemical fertilizer, insecticides, herbicides, and improved or hybrid seed. On farms where farmers are unaccustomed to making purchases with cash, great care must be taken to evaluate the productivity of, or return to, the additional amount of cash required for alternative technologies.

On fully commercialized farms, where cash is basically not a limiting factor, the criterion of profit maximization may be relevant. Profit maximization is achieved where the value of the product obtained from the last unit of input is just equal to the cost of that additional unit. However, farmers with very limited amounts of cash will not usually be interested in utilizing as much cash in an individual enterprise as is required to maximize profit. Rather, they will be looking for ways to achieve the highest return (or productivity) per unit of cash invested in the enterprise. In this situation, the amount of product per unit of cash (similar to the amount of product per unit of labor) is a relevant evaluation criterion. Calculations can be made in a manner similar to those shown above for returns to labor.

Because cash can be converted into many different kinds of inputs, it is more critical to look at alternative uses for cash and not just consider return to cash investment for individual enterprises. This is even more critical on small farms where family necessities compete directly for limited cash resources. If researchers consider only the return to cash investment in the commodity in which they are interested, they may well find that what appears to be a "good" technology is not acceptable to farmers, who would rather use the cash in another way, such as for a wedding or to repair the house.

RESPONSE SURFACES*

What They Are

A response surface is a representation of the natural relationship which exists between the quantity of a product and various levels of one or more inputs used to produce that product. That is, a response surface is an estimate of the response of a product to different quantities of one or more inputs utilized in the production process. This representation can be physical, tabular, graphic, or mathematical. A true response surface with three dimensions (width, depth, and height) results in graphic form when only two inputs are used to obtain one product. In this form, the surface is similar to a hill with quantities of the two inputs measured toward the north and east and quantity of product represented by height, or altitude. Each point on the surface represents a combination of different quantities of input A, input B, and yield of the product. In the simplest case of only one input, the "surface" is a straight line or a curve. In a more complex case with three or more inputs used to produce one product, it is impossible to imagine or draw the response "surface," so it must be represented in mathematical form.

The graphic or mathematical form of a response surface is an artificial estimate of a real phenomenon that exists in nature. The surface can represent the response of a product to a continuous input, such as a crop to fertilizer or animals to feed, etc. Because the response is a natural biological phenomenon, it is not constant for a fixed level of inputs. Rather, it is subjected to a random variance. The graphic or mathematical response surface represents the central tendency of the response in terms of magnitude (level) and of form (curvature).

Representative Forms

To demonstrate the forms which response surfaces can take, a two-dimensional case will be used with one input

*Much of this section was translated and adapted from Hildebrand (1972).

and one product. Surfaces can be represented either graphically or mathematically in generalized form. Mathematically, yield (Y) is a function of the quantity of input (X), that is Y = f (X).

The simplest form of a response surface, one that is seldom useful, is a straight line (Fig. VI-1a). Generally, linear responses are found only with small changes in the level of the input.

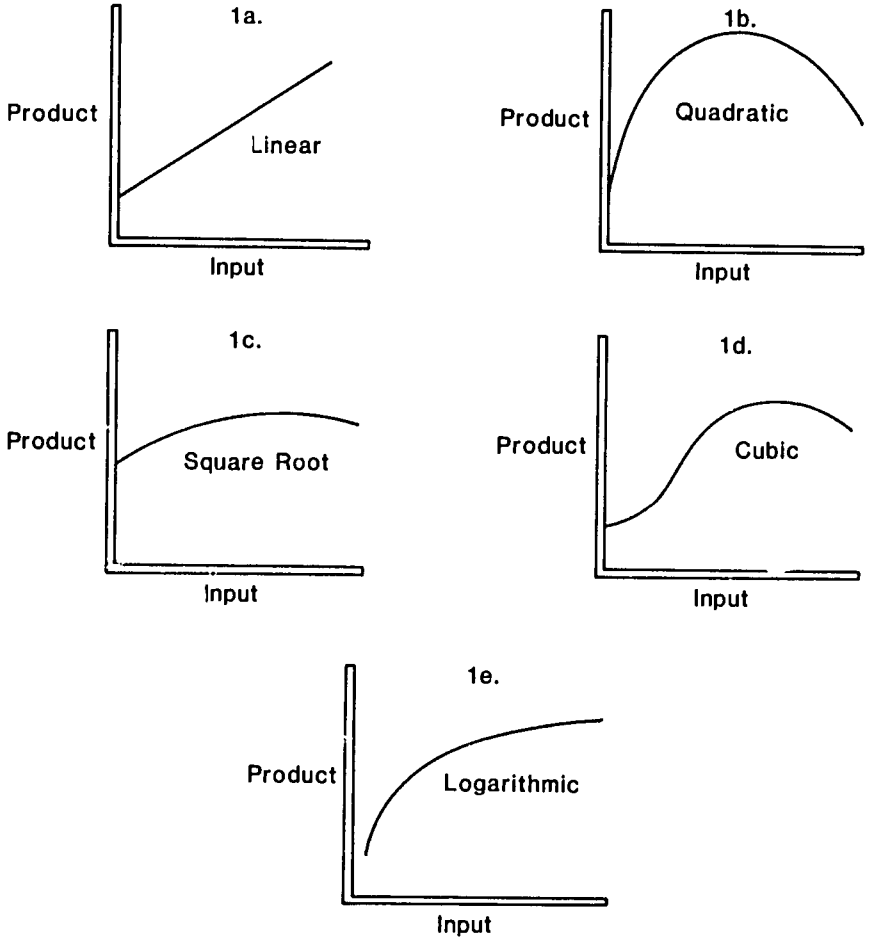


FIG.VI-1. Two-dimensional response surfaces.

Perhaps a more useful form of curve is the parabola, or quadratic function (Fig. VI-1b). This type of curve is representative of part of the law of diminishing returns. It has a maximum and can demonstrate a decline in production resulting from high or toxic use of an input if that phenomenon is present. A similar form is the square root function (Fig. VI-1c). The square root function also has a maximum but it is less sharp than the parabola. In some cases, data represent surfaces which are more complex than can be represented by the previous curves. If the response surface contains all three economic stages of production, a form such as a cubic function (Fig. VI-1d) is required. A cubic function can have a portion that increases at an increasing rate (Stage I), another part that increases at a decreasing rate (mostly in Stage II), and finally a portion that decreases with increases in the level of the input (Stage III). Another useful function is the logarithmic function (Fig. VI-1e), called by economists the Cobb-Douglas function. This function does not have a maximum and is therefore useful for responses that display this characteristic (for example, herbicides and insecticides if they do not induce decreasing production). It is also an easier function to work with than are some other forms.

Economic Analysis

The analysis of data by response surfaces is related to analysis of variance, but it is more efficient in describing the relationships and results in more directly applicable information for researchers and farmers. The least-squares method, which is used to calculate the statistical values of the surfaces by regression, is based on all usable observations and is therefore efficient in the use of the data. Also, once a researcher has a mathematical function, it is possible to interpolate and predict responses for input levels not included in the original experimental design. With care, estimated responses can be extrapolated beyond the range of the data. Also, with a curvilinear function it is possible to find the quantity or the combination of inputs which result in maximum physical production by equating the first derivative of the response surface to zero.

Perhaps the most useful reason to calculate response

surfaces is to facilitate economic analysis. After researchers have calculated the equation for the response surface, they can calculate the quantity or combination of inputs which are most profitable for commercial producers from any combination of prices of inputs and product. The economic optimum combination of inputs is the amount which results in maximum profit for farmers (Fig. VI-2). Using U for profit (utility), which is the difference between income (I) and cost (C) for one input, the following calculations can be made:

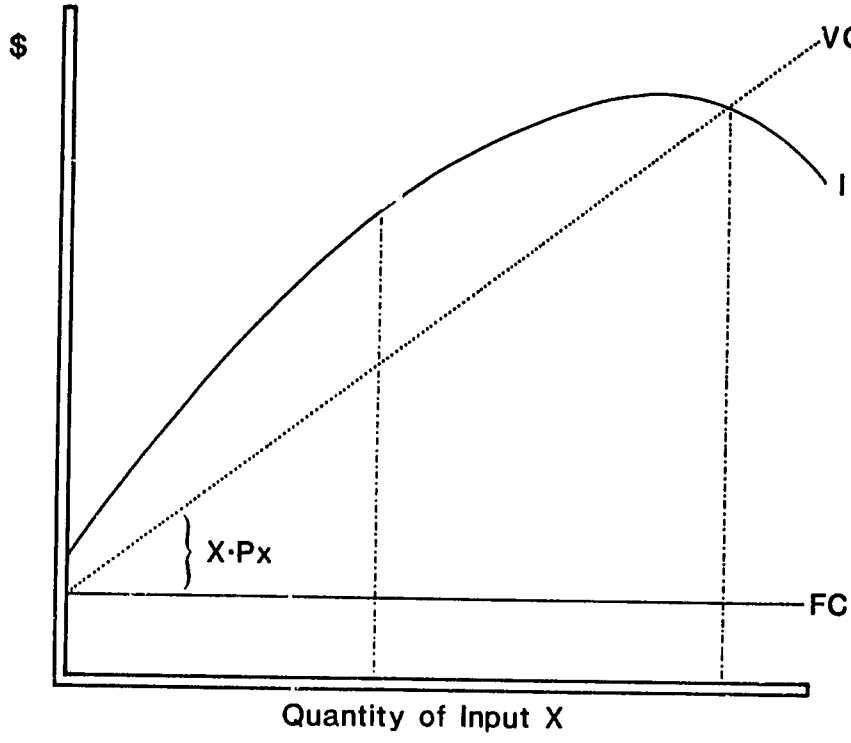
$$I = Y(P_y)$$
$$C = X(P_x) + FC$$
$$U = I - C$$
$$= Y(P_y) - X(P_x) - FC$$

where:

I = total income
Y = quantity of product (generally yield per ha) and is a function of X: $Y = f(X)$
C = total cost (variable + fixed cost)
P_y = price of the product
X = quantity of input
P_x = price of the input
FC = fixed costs that include all costs except the purchase of X, which is the input under consideration
U = profit or net income (utility)

In the rational or economic stage of production, the response curve is increasing, but at a decreasing rate. Therefore, the function of I (total income) will have the same form. Cost is a straight line with an intercept equal to FC. The difference between I and C, which is utility or profit, is a curve which has a maximum at the point where the vertical difference between I and C is greatest. The quantity of the input that results in this value is the quantity which maximizes profit. Mathematically, this can be calculated as follows:

$$U = Y(P_y) - X(P_x) - FC$$
$$dU/dX = [(dY/dX)\ P_y] - P_x$$
$$= \text{(marginal profit)}$$



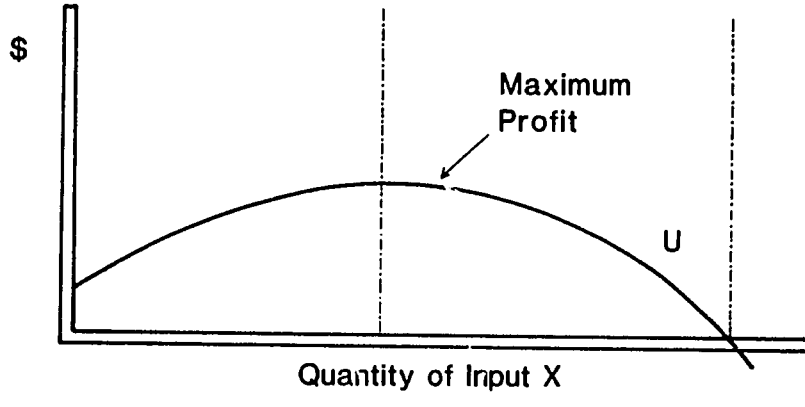


FIG.VI-2. Variable cost (VC), fixed cost (FC), income (I), and profit (U) functions.

When marginal profit is equal to zero, the profit curve U is at a maximum. Therefore, in order to find the amount of the input which results in maximum profit, the following relationship is calculated:

$$[(dY/dX)\ P_y] - P_x = 0\ , \quad \text{or}$$
$$dY/dX = P_x/P_y$$

In other words, to maximize profit, the derivative of the response surface (dY/dX) is equated to a ratio of the price of the input to the price of the product (Px/Py). The prices of the inputs and the prices of the products must be measured in the same units used in calculating regression.

For two or more inputs, the solution is an extension of the above:

$$U = Y(P_y) - X_1(Px_1) - X_2(Px_2) - \ldots - X_n(Px_n) - FC$$

where $X_1, X_2, \ldots, X_n$ are the inputs included as variables in the response surface. The partial derivatives (these are partial derivatives because there is now more than one input, X) are:

$$\delta U/\delta X_1 = [(\delta Y/\delta X_1)\ Py] - Px_1 = 0$$
$$\delta U/\delta X_2 = [(\delta Y/\delta X_2)\ Py] - Px_2 = 0$$
$$\vdots$$
$$\delta U/\delta X_n = [(\delta Y/\delta X_n)\ Py] - Px_n = 0$$

The simultaneous solution of this set of partial derivatives will result in the combination of the n inputs to produce product Y that maximizes profit for farmers.

There are various ways to estimate response surfaces. The best in any case will depend on the quantity of inputs, the number of treatments, and the type of calculating or computing equipment available to researchers.

METHODS OF OBTAINING RESPONSE SURFACES

Simple Regression--Visiographic

If an experimental design includes only one input and few treatments, and the researchers desire a quadratic form of response surface, a simple and practical method is to estimate the curve visually and graphically. This requires only the simplest type of calculator. A design using three levels of an input provides an exact estimate of a quadratic curve. With more than three levels, the method is less exact but still functional. Table VI-1 shows partial data from a fertilizer experiment on rice conducted in Colombia.

Table VI-1. Nitrogen effect on yield of IR-22 rice in Corinto, Cauca, Colombia.

Treatments			Yields		
			Replications		
N	P_2O_5	K_2O	I	II	Average
			kg/ha		
0	40	40	5500	5200	5350
150	40	40	7000	6350	6675
300	40	40	6950	7000	6975

Source: Hildebrand, 1972

Nitrogen is the only input that varies in these data. The yield of the individual replications and the mean of each treatment level are shown graphically in Fig. VI-3. The researchers desire to estimate a quadratic function of the form

$$\hat{Y} = a + b_1X + b_2X^2$$

which represents the response of rice to nitrogen under these conditions.

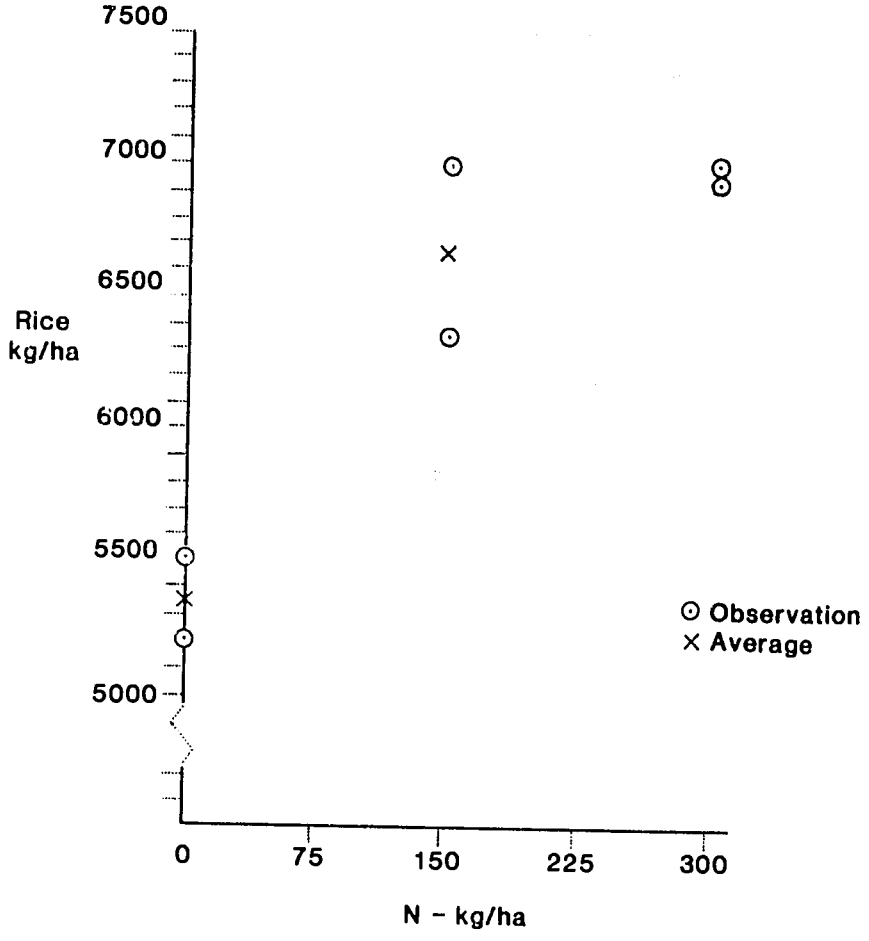


FIG.VI-3. Response of IR-22 rice to nitrogen in Corinto, Cauca, with P_2O_5 = K_2O = 40 kg/ha.

The method is based on the fact that the first derivative of a quadratic function is a straight line (Fig. VI-4). The first derivative is a value equal to the slope of the function for any value of X, or in this case the input nitrogen. Associated with an increase in nitrogen from 0 to 150 kg/ha (ΔN_1) is an average increase of 1325 kg in the yield of rice (ΔY_1). Here, the symbol Δ means "change in." Hence, ΔN_1 can be read, "the change

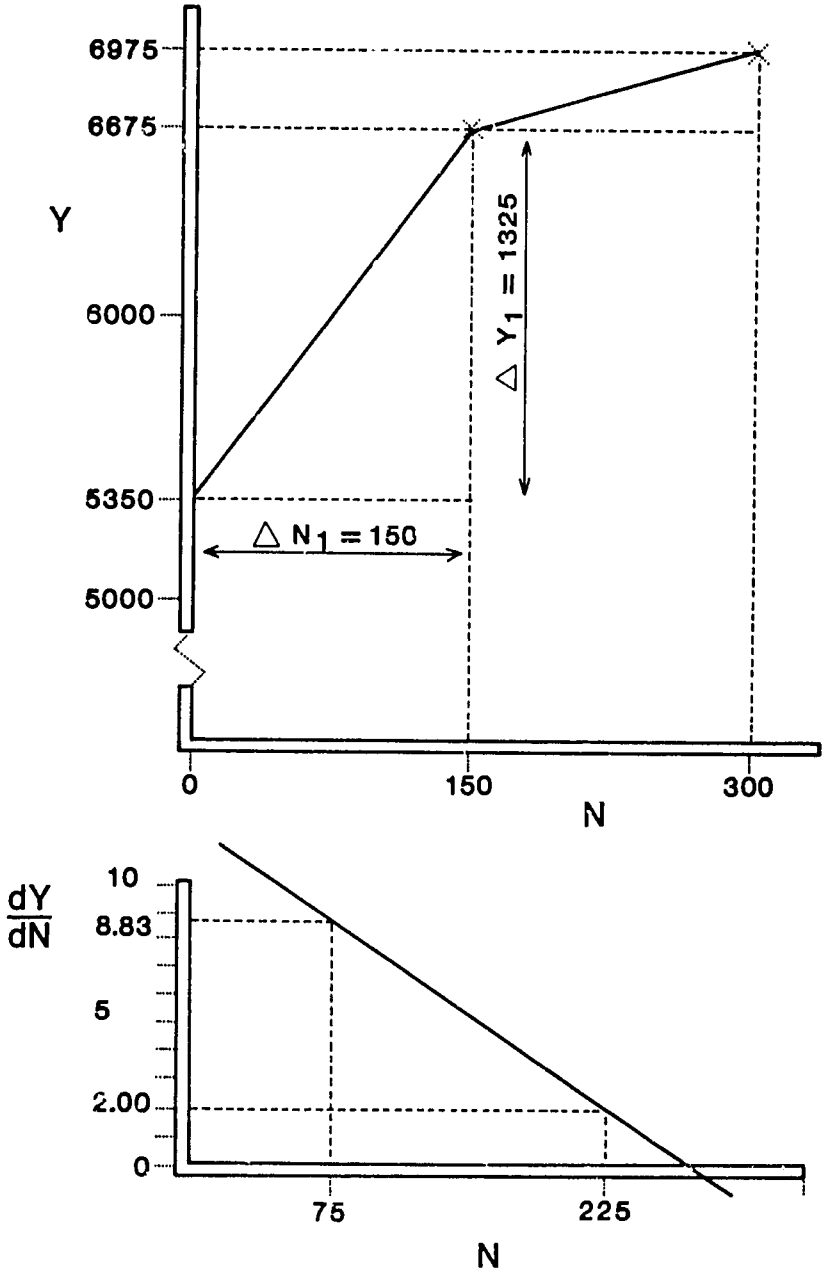


FIG.VI-4. Method for estimating a quadratic response surface.

in N_1." The ratio $\Delta Y_1 / \Delta N_1$ is equal to the slope of the line which connects the mean yield values for 0 and for 150 kg/ha of nitrogen. This is also the average slope of the curve between 0 and 150 kg/ha of nitrogen or equivalent to 75 kg of nitrogen (Fig. VI-4). Hence, the value of $\Delta Y_1 / \Delta N_1 = 8.83$ is plotted on the lower portion of Fig. VI-4 for N = 75. The slope of the response surface between 150 and 300 kg/ha of nitrogen is calculated the same way. Here, ΔY_2 is 300 and ΔN_2 is 150, so $\Delta Y_2 / \Delta N_2 = 2$. This value is plotted midway between 150 and 300 or equivalent to N = 225. The straight line drawn between the two values of $\Delta Y/\Delta N$ plotted on the lower part of Fig. VI-4 is the function of the slope, or the first derivative of the quadratic response surface.

Having two points on a straight line, it is simple to calculate the equation of this line. The slope is derived in the same manner as the slope for the curve. Find the change in value of the derivative and divide that by the change in value of N. In this case there are two known points. For N = 75, the value of $\Delta Y/\Delta N$ is 8.83 and for N = 225, the value of $\Delta Y/\Delta N$ is 2. Hence, for the straight line the value of $\Delta(\Delta Y/\Delta N) = 8.83 - 2.00 = 6.83$. The difference between 225 and 75 is 150, so the slope of the straight line is: $\Delta(\Delta Y/\Delta N)/\Delta N = -6.83 / 150 = -0.0455$. It is negative because the value $\Delta Y/\Delta N$ diminishes between the values of N = 75 and N = 225.

To this point, then, a portion of the equation of the straight line is known. That is, the value of the slope b equals -0.0455, so Y = a - 0.0455 N. To find the value of a, known values of $\Delta Y/\Delta N$ and N can be substituted to solve the equation for a. For N = 75, the value of $\Delta Y/\Delta N = 8.83$. Substituting these in the above equation, produces $-8.83 = a - 0.0455 (75)$. Solving this equation for a gives a = 12.25. The equation of the derivative of the quadratic curve, then, is

$$\Delta Y/\Delta N = 12.25 - 0.0455 N$$

Mathematically, the quadratic response curve is the integral of the straight line equation above. A simple means of finding the integral of a straight line is to multiply the equation by N (in this case) and divide the slope of the straight line by 2. The result,

$12.25\ N - 0.0228\ N^2$, is part of the quadratic function. To this a constant of integration must be added, which is equivalent to the intercept or the value of Y when N is zero. When a zero-level treatment is included in the data set, and there are only three levels of input, the average yield at the zero level can be used as C, the constant of integration. In this case the average yield of the zero treatment for IR-22 rice in the trial is 5350 kg/ha, so this becomes the value of C. The complete quadratic equation, then, is:

$$\hat{Y} = 5350 + 12.25\ N - 0.0228\ N^2$$

Later, using more orthodox statistical procedures, we will show that this is the same equation that can be derived by the method of least squares.

If no zero treatment is included in the data set, it is still possible to find the value of C, the constant of integration. This is done once again by using known values of Y and N and solving for the unknown constant of integration. For example, in this case, for a value of N = 150 the average value of Y = 6675 kg/ha. Hence, the following information is available:

$$6675 = C + 12.25\ (150) - 0.0228\ (22{,}500)$$

which comes from the values of N = 150 and N^2 = 22,500. Solving this equation for C results in C = 5350.5, which is essentially the same value that was calculated before. Another alternative is to use the mean value of N and the mean value of Y. However, care must be taken to calculate the mean value of N^2 from the original levels of N. For example, the three levels of N are 0, 150, and 300. Squaring each of these values gives 0; 22,500; and 90,000. The average of these three numbers is 37,500.

The quadratic response surface being used to represent the data is shown in Fig. VI-5. Note that in order to make the curve smooth, the equation can be solved for levels other than those used in the experiment. For example, a level of 50 kg/ha of N results in an estimated value of Y for rice of 5905.

Now that the response surface has been estimated, it can be used to help analyze the data. First, find the level of nitrogen which maximizes physical production. As

stated previously, this value of nitrogen can be found by taking the derivative of the response surface and equating it to zero. The derivative of the quadratic function is

$$d\hat{Y}/dN = 12.25 - 0.0455\ N$$

This equation is then set equal to zero and solved for N. In this case, N = 270 kg/ha, which will result in a

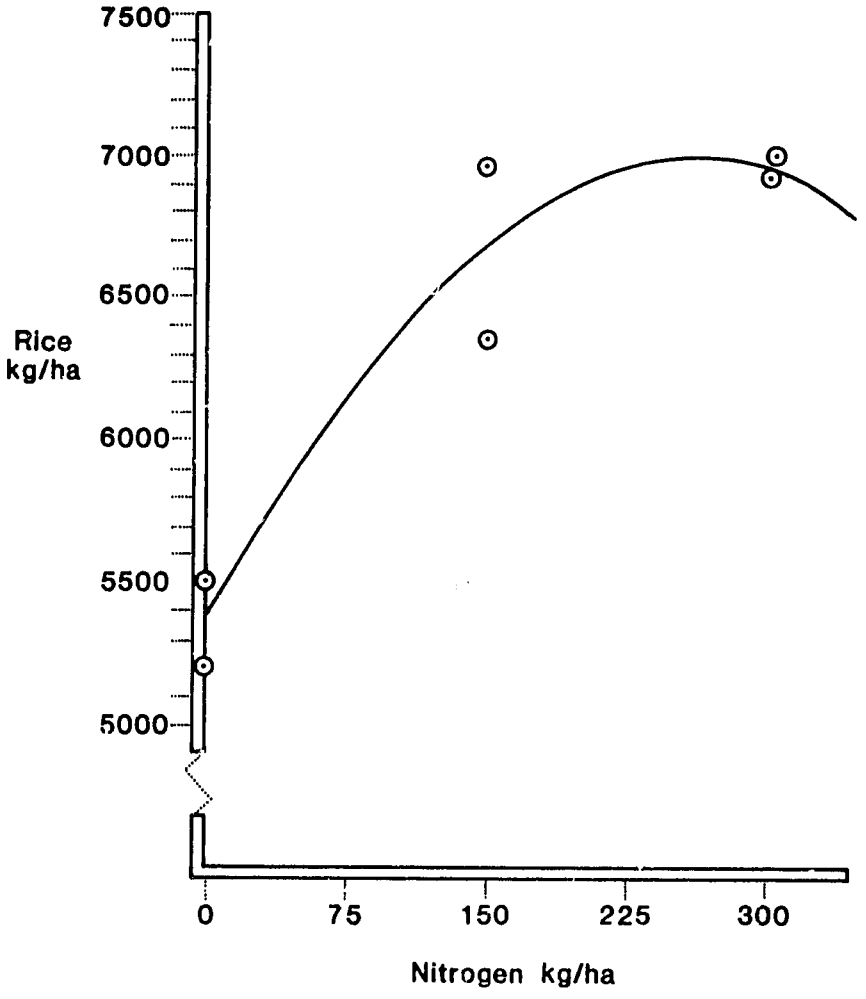


FIG.VI-5. Estimated quadratic surface.

maximum yield per hectare of rice. Solving the equation of the response surface for 270 kg of N shows that maximum expected rice production would be 6995 kg/ha. This point can be plotted on the curve.

Economically, a more important value would be the amount of nitrogen which would maximize profit for the farmer with abundant resources. Once again, referring to a previous discussion, this value can be found by equating the derivative of the response surface to the ratio: price of nitrogen/price of rice. Remember that the prices must be in the same units of measurement as the equation. At the time of the study, IR-22 rice had a value of $2.80/kg and the price of nitrogen as urea was $4.40/kg, both expressed in Colombian pesos. Therefore, the following relationship can be found:

$$d\hat{Y}/dN = 12.25 - 0.0455\ N = 4.40\ /\ 2.80,$$
$$\text{from which}\quad N = 235$$

In other words, 235 kg/ha of N will maximize profit for farmers with the prices as stated. Once again, solving the equation of the response surface for N = 235, the value of Y = 6970 can be found. Although it is possible to produce 6995 kg/ha of rice with 270 kg of N, it is not profitable to apply more than 235 kg/ha, because the cost of the additional 35 kg/ha of N is greater than the value of the additional 25 kg of rice.

It is possible to use the visiographic method to calculate response surfaces for more than three levels of one input, but the results are less precise. It is necessary to draw a curve through the data points visually in a form that approximates a quadratic curve and appears to be representative of the data. Once the shape of the curve satisfies the researcher, three points along that curve can be chosen and the same calculations made as above. When the equation of the estimated response surface is found, if it does not satisfy the researcher, it is possible to begin again with a new curve drawn visually. Obviously, this method is subject to human error and the bias of the researcher. However, with practice it is possible to achieve acceptable results (see Fox, 1968, p.102.) even though no estimates of sampling error will exist.

Simple Regression--Least Squares

With hand-held calculators, simple regression can be figured easily. However, it is worthwhile to understand the procedure in case such a calculator is not available.

Table VI-2. Data and calculations for quadratic regression.

	X_1	X_2	Y
	0.0	0.0	5.50
	0.0	0.0	5.20
	1.5	2.25	7.00
	1.5	2.25	6.35
	3.0	9.0	6.95
	3.0	9.0	7.00
Total	9.0	22.50	38.00
$\bar{X}$ or $\bar{Y}$	1.5	3.75	6.333

Note: $X_1 = N$ and $X_2 = N^2$

In order to compare the statistical method with the visiographic method, the same data set will be used. To facilitate the calculation by a calculator, the data will be tabulated in a somewhat different form (Table VI-2). To avoid the use of very large numbers, the data have been coded. An appropriate coding procedure is to consider N in units of 100 kg and yield of rice in metric tons. In this way, all the values fall between 0 and 10. The required calculations follow. Take careful note of capital and lower-case letters.

$$\Sigma x_1^2 = \Sigma X_1^2 - [(\Sigma X_1)^2 / n]$$

where $\Sigma X_1^2 = (1.5)^2 + (1.5)^2 + (3.0)^2 + (3.0)^2$

and $(\Sigma X_1)^2 / n = 9^2 / 6$
$= 13.5$

then $\Sigma x_1^2 = 22.5 - 13.5$
$= 9.0$

$$\Sigma x_2^2 = \Sigma X_2^2 - [(\Sigma X_2)^2 / n]$$
$$= 172.125 - 84.375$$
$$= 87.750$$

$$\Sigma x_1 x_2 = \Sigma X_1 X_2 - [(\Sigma X_1)(\Sigma X_2) / n]$$
$$= 60.75 - 33.75$$
$$= 27.00$$

$$\Sigma x_1 y = \Sigma X_1 Y - [(\Sigma X_1)(\Sigma Y) / n]$$
$$= 61.875 - 57.000$$
$$= 4.875$$

$$\Sigma x_2 y = \Sigma X_2 Y - [(\Sigma X_2)(\Sigma Y) / n]$$
$$= 155.5875 - 142.5000$$
$$= 13.0875$$

$$\Sigma y^2 = \Sigma Y^2 - [(\Sigma Y)^2 / n]$$
$$= 243.9150 - 240.667$$
$$= 3.2483$$

The desired function has the general form

$$\hat{Y} = a + b_1 X_1 + b_2 X_2$$

The values of a, b_1, and b_2 must be found. First, find the value that is normally called D:

$$D = (\Sigma x_1)(\Sigma x_2^2) - (\Sigma x_1 x_2)^2$$
$$= (9.0)(87.75) - (27)^2$$
$$= 60.75$$

Then calculate b_1 by the following equation:

$$b_1 = [(\Sigma x_2^2)(\Sigma x_1 y) - (\Sigma x_1 x_2)(\Sigma x_2 y)] / D$$

and b_2 as follows:

$$b_2 = [(\Sigma x_1^2)(\Sigma x_2 y) - (\Sigma x_1 x_2)(\Sigma x_1 y)] / D$$

In the example:

$$b_1 = [(87.750)(4.875) - (27.00)(13.0875)] / 60.75$$
$$= 1.225$$

$$b_2 = -13.8375 / 60.75$$
$$= -0.228$$

Now that b_1 and b_2 are known, find the intercept a by using the mean values of Y, X_1, and X_2 as follows:

$$6.333 = a + 1.225(1.5) - 0.228(3.75)$$

from which

$$a = 5.350$$

Remembering that X_1 is N and X_2 is N^2, the complete equation is:

$$\hat{Y} = 5.350 + 1.225\ N - 0.228\ N^2$$

This is the same equation as was found by the visiographic method, using the average values of the replications. The values in this equation have been coded so the magnitude of the numbers is different. Here N is measured in units of 100 kg and Y in metric tons. Maximum production results from N = 2.7 units (270 kg), and production at this level is 6.99 tons (6990 kg). These are essentially the same values found before.

In order to find the level of N which maximizes profit, use the prices for the same units in which N and rice are measured. Because 1 kg of N costs $4.40, the price of 100 kgs is $440. With rice at $2.80/kg the price of a ton is $2800. In order to find the amount of N which will maximize profit, calculate:

$$dY/dN = 1.225 - 0.456\ N = 440/2800$$
so N = 2.34 units or 234 kg

This is essentially the same value as was found by the visiographic method. The difference is due to slight roundoff error from coding.

ANALYSIS OF RESPONSE SURFACES

Statistical tests of response surfaces are based on three measures of variations in the dependent variable (frequently yield): 1) the variation explained by regression, 2) unexplained variation, and 3) total variation or the sum of 1) and 2). A surface relating the response of rice to N can be described as:

$$Y = f(N|P, K, \text{water, soils, management, etc.}),$$

where N is the variable input and those factors to the right of the vertical line | are fixed or non-variable. In the generalized function, there are terms relating the yield to quantities of N and there is an error term:

$$\hat{Y} = a + b_1 N + b_2 N^2 + e$$

Included in the error term (e) are errors in measurement of the quantity of N applied (although this is assumed to be precise); error in controlling levels of P, K, irrigation, etc.; differences in management; and other factors, such as incidence of pests that affect some plots more than others. Also included in the error term is the error which results because the form of the curve utilized does not perfectly fit the curvature of the biological phenomenon. In summary, explained variation is comprised of $a + b_1N + b_2N^2$, and unexplained variation is contained in the term e.

Total variation in yield is measured by the sum of the squared deviations of yield from the average yield, designated as y^2. The deviations from the mean value of Y are shown in Fig. VI-6a. Explained variation (variation due to regression) can be shown as:

$$\Sigma\hat{y}_{12}^2 = \Sigma\{b_i(\Sigma x_i y)\}$$

Unexplained variation (the sum of squared deviations from the regression) is:

$$\Sigma(Y - \hat{Y})^2 = \Sigma y^2 - \Sigma\hat{y}_{12}^2$$

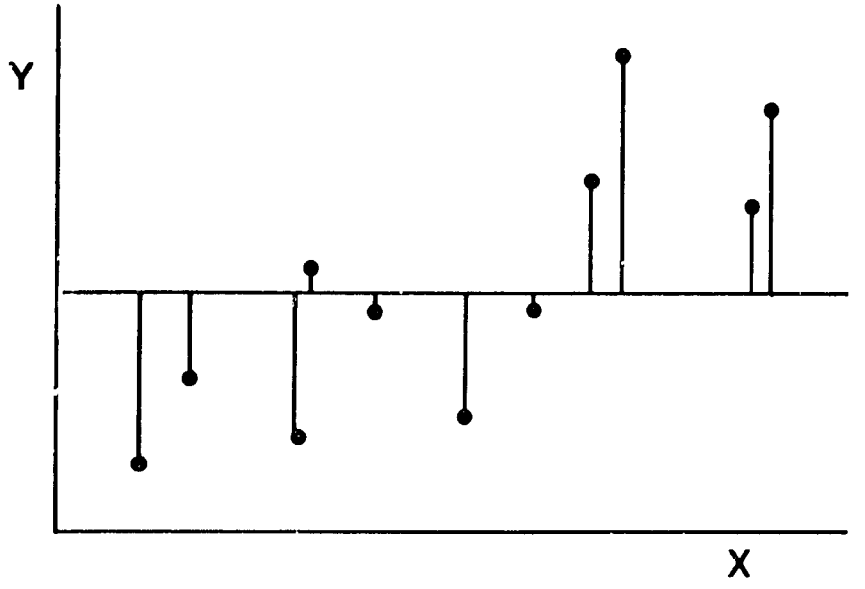


FIG.VI-6a. Deviations from the mean value of Y.

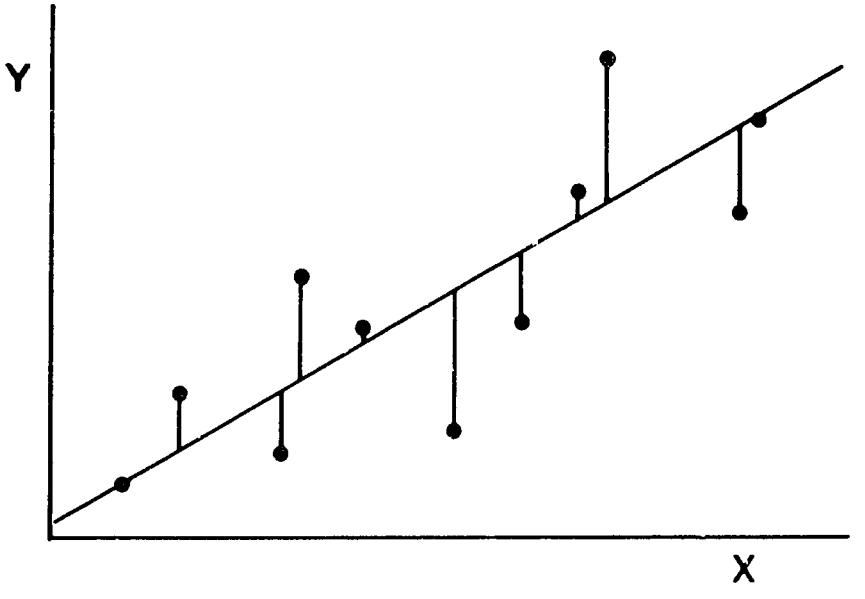


FIG.VI-6b. Deviations from regression.

Deviations from regression are shown in Fig. VI-6b. Measures for evaluating the regression will be based upon the rice data previously used for the visiographic estimates.

Coefficient of Determination

One of the best-known statistical measures of how well a regression fits the data is the coefficient of determination or the "goodness of fit," R^2. This coefficient is the proportion or percentage of the total variation which is explained: explained variation divided by total variation. Explained variation for the quadratic function is:

$$\Sigma\hat{y}_{12}^2 = b_1\Sigma x_1 y + b_2\Sigma x_2 y$$

or in the example:

$$\begin{aligned}\Sigma\hat{y}_{12}^2 &= 1.225(4.875) - 0.228(13.0875)\\ &= 2.9879\end{aligned}$$

Total variation is

$$\Sigma y^2 = 3.2483$$

Therefore,

$$\begin{aligned}R^2 &= (b_1\Sigma x_1 y + b_2\Sigma x_2 y) \;/\; \Sigma y^2\\ &= 2.9879 \;/\; 3.2483\\ &= 0.920\end{aligned}$$

This value of 0.920 means that 92% of the variation in yield, when nitrogen is added, is explained by the calculated quadratic response surface. As a general (but not firm) rule, the value of R^2 should be greater than 0.5. In other words the response surface calculated should explain at least half the variation in the data. An R^2 value of 0.92 is considered high for the kind of data being used here.

Variance, Standard Error, and Confidence Intervals

The standard error of the estimate of the regression is the square root of the unexplained variation adjusted for degrees of freedom. Unexplained variation, or the variance of the regression, is

$$S_{y.12}^2 = (\Sigma y^2 - \Sigma \hat{y}_{12}^2) / (n - k - 1)$$

where k is the number of coefficients estimated (b_1 and b_2 in this example). In the example:

$$S_{y.12}^2 = (3.2483 - 2.9879) / (6-2-1)$$
$$= 0.0868$$

Standard error is the square root of the variance:

$$S_{y.12} = (S_{y.12}^2)^{1/2}$$
$$= (0.0868)^{1/2}$$
$$= 0.29462$$

This value indicates the precision with which the regression measures the average value of yield, $\bar{Y}$ = 6.333 tons/ha. This value is used to calculate the confidence interval around the mean yield as follows:

$$\bar{Y} \pm t_a \, [S_{y.12} / n^{1/2}]$$

where t_a is the tabulated value of t for the level of significance a and with n - k degrees of freedom.

For example, at the 95% confidence level (a = 0.05) with n = 6 and with 6 - 2 = 4 degrees of freedom, the confidence interval is:

$$6.333 \pm (2.776)(0.29462 / 2.45)$$

or

$$6.333 \pm 0.334$$

With this result, there is a 95% confidence that the true mean yield for the same treatments is included between 5.999 and 6.667 metric tons/ha.

Significance of the Coefficients

The confidence interval about the mean is a way to measure the error associated with the level of the response surface. It is also possible to have errors in measuring the slope and the curvature of the response surface, represented by the values of b_1 and b_2 in the example. The usual test is to determine whether the values of the b's (the regression coefficients) individually are different from zero. This is done by the use of the "t" test. If the value of t is not sufficient to provide confidence that the value of a certain b_i is different from zero (either positively or negatively) it is concluded that yield is not affected by the term corresponding to the b_i being tested. This follows, because if the value of b_i is not different from zero, then it must equal zero and the term disappears from the equation. Hence, in the equation

$$\hat{Y} = 5.350 + 1.225\ N - 0.228\ N^2$$

if it is determined that b_2 (-0.228) is not different from zero, the N^2 will disappear because it is multiplied by zero and the equation left is of the form $Y = a + bN$. The value of b in the new equation would have to be calculated again because it would be different from b_1 in the quadratic form.

The t values for each of the coefficients are calculated as follows:

$$\begin{aligned} t_1 &= b_1 / [(S_{y.12})(\Sigma x_2^2 / D)^{1/2}] \\ &= 1.225 / [(0.29462)(87.75 / 60.75)^{1/2}] \\ &= 3.459 \end{aligned}$$

$$\begin{aligned} t_2 &= b_2 / [(S_{y.12})(\Sigma x_2^2 / D)^{1/2}] \\ &= -0.228 / [(0.29462)(9.0 / 60.75)^{1/2}] \\ &= -2.0160 \end{aligned}$$

Because it is known that the value of b_1 should be positive and the value of b_2 should be negative, it is possible to use a one-tailed "t" table for testing the level of significance of these coefficients. If one does not know beforehand whether a coefficient should be positive or negative, it is necessary to use a two-tailed

table because the coefficient could be either greater or less than zero. Frequently, a two-tailed table is required to test the value of interaction terms, for example. Note that the most commonly available "t" table is the two-tailed table. Returning to the example, find that for $n - k - 1 = 6 - 2 - 1 = 3$ degrees of freedom, the value of t in a one-tailed table for the 0.025 level of confidence is 3.182. The value of t for the 0.05 level of confidence is 2.353 and for the 0.10 level of confidence the value of t is 1.638. From this, it can be seen that the coefficient b_1 is significant at the 2.5% level and b_2 is significant at the 10% level. These values of t (one-tailed) can also be taken from a two-tailed table, as follows:

Distribution of t for 3 degrees of freedom

	Probability of a larger value		
One-tailed	0.100	0.050	0.025
Two-tailed	0.200	0.100	0.050
t	1.638	2.353	3.182

The confidence interval for the two b_i coefficients is a function of t_a and $S_{y.12}$. Obviously, different values of the b_i within the confidence interval will create different kinds of response surfaces. Errors in estimating the b_i coefficients create wider errors as one moves further from the mean value. Hence, the confidence band (Fig. VI-7) is narrowest near the mean values of X and Y. Errors in estimating the coefficients plus error in estimating the level of the regression create the confidence band for the regression.

Significance of the Equation

The t test determines the significance of each of the coefficients individually. Each coefficient can be accepted or rejected in accordance with results of the test. The F test is a test of the significance of the complete equation and, at the same time, a test of all the coefficients combined. The value of F is calculated from already-known values of explained variation and the variance of the regression, as follows:

$$F_c = (\Sigma\hat{y}_{12}^2 / k) / S_{y.12}^2$$
$$= (2.9879 / 2) / 0.0868$$
$$= 17.2114$$

The F value in the table for (k) and (n - k - 1) or 2 and 3 degrees of freedom is 16.04 for the 0.025 level of probability. This means that the response surface being calculated is highly significant (the calculated value of F is larger than the value in the table).

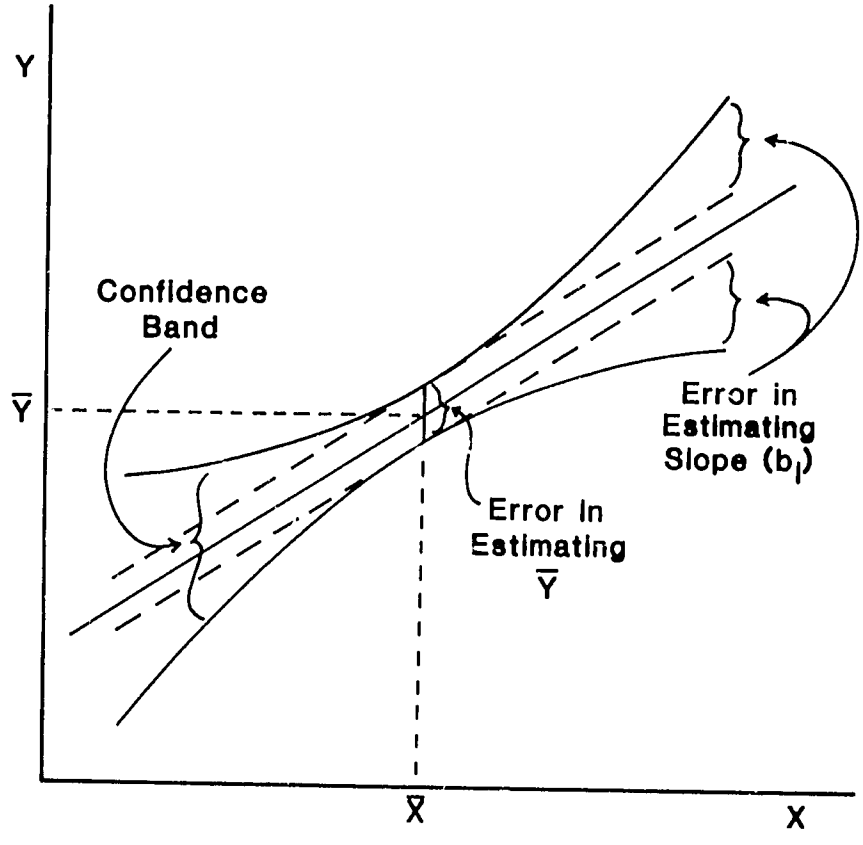


FIG.VI-7. Errors in estimation contributing to a confidence band in regression.

In partial summary, the following values have been calculated to help determine the significance of the equation:

$R^2 = 0.920$
$t_1 = 3.459$ (2.5%)
$t_2 = 2.0160$ (10%)
$F_c = 17.2114$ (2.5%)

In general, these values and levels of significance are high, which indicates that the response surface calculated is very representative of the real biological response demonstrated by the data. For this reason, it is probably not necessary to make other calculations to study the significance of the response surface. Nevertheless, in order to demonstrate the use of other analyses, the same equation will be used.

Linear versus Quadratic Surface

In some cases, it may be logical to wonder if a linear response better fits the data than a curvilinear response. From the same rice data (Table VI-2), a linear equation of the form $Y = a + bN$ can be calculated. The value of b is:

$$b = \Sigma x_1 y \,/\, \Sigma x_1^2$$
$$= 4.875/9.0$$
$$= 0.5416$$

Using the mean value of Y and X, the value of <u>a</u> is:

$$a = \bar{Y} - b(\bar{X}_1)$$
$$= 6.333 - 0.5416(1.5)$$
$$= 5.5206$$

The linear equation is:

$$\hat{Y} = 5.5206 + 0.5416\ N$$

Note that the b in the linear equation is different from the value of b_1 in the quadratic equation.

The "goodness of fit" of the linear equation can be evaluated by

$$R^2 = b\Sigma x_1 y \;/\; \Sigma y^2$$
$$= 0.5416(4.875) \;/\; 3.2483$$
$$= 0.81$$

which is lower than for the quadratic equation.

The significance of b can be tested using

$$t_b = \frac{b}{\{(\Sigma y^2 - [(\Sigma x_1 y)^2 \;/\; \Sigma x_1^2]\}^{1/2} \;/\; [(n-2)\Sigma x_1^2]^{1/2}}$$

$$= 0.5416 \;/\; \{(3.2483-[(4.875)^2 \;/\; 9]) \;/\; 4(9)\}^{1/2}$$
$$= 4.1686$$

and with $n - k - 1 = 4$ degrees of freedom, b is significantly different from zero.

The F value for the linear equation is:

$$F_c = t_b^2$$
$$= 17.38$$

With k and n - k - 1, or 1 and 4, degrees of freedom this is significant at the 5% level.

Another test of the linear versus the quadratic equation is the "F" test, based on an analysis of variance. The method is used for evaluating the significance of additional terms. Total variation (Σy^2) is divided into three parts: (1) the variation explained by the equation without the additional term, that is, the linear equation; (2) the increase in the explained variation resulting from adding the additional (quadratic) term; and (3) the error or residual. The part explained by the equation without the additional term is

$$\Sigma(b_i \Sigma x_i y)$$

where i varies from 1 to k - 1, and the increase in explained variation from the additional term is

$$\Sigma(b_j \Sigma x_j y) - \Sigma(b_i x_i y)$$

where j varies from 1 to k and i varies from 1 to k - 1 as before. The term k is the additional term. Because different terms are included in the two equations, the

values of b_i will be different in each equation. For this reason, the increase obtained from the additional term is <u>not</u> simply $b_k \Sigma x_k y$.

The value of the sources of variation for the ANOVA (Table VI-3) in the rice example are calculated as follows:

Linear: $b\Sigma x_1 y = 0.5416(4.875)$
$= 2.603$

Addition of quadratic term:

$b_1 \Sigma x_1 y + b_2 \Sigma x_2 y - b\Sigma x_1 y$
$= 1.225(4.875) - 0.228(13.0875) - 0.5416(4.875)$
$= 5.972 - 2.984 - 2.6403$
$= 0.3477$

Total $\Sigma y^2 = 3.2483$

Error: total - linear - quadratic
$= 3.2483 - 2.6403 - 0.3477$
$= 0.2603$

TABLE VI-3. Analysis of variance comparing linear and quadratic equations.

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares
Linear equation	k-1 = 1	2.6403	
Addition of the quadratic term	1	0.3477	0.3477
Error or variation of the quadratic equation	n-k-1 = 3	0.2603	0.0868
Total	n-1 = 5	3.2483	

The mean square values in Table VI-3 are obtained in the usual way, by dividing the sum of squares by the

appropriate degrees of freedom. The value of F_c is calculated by dividing the mean square of the additional term by the mean square error or

$$F_c = 0.3477 / 0.0868$$
$$= 4.0058$$

d.f. = (1) and (n-k-1) or 1 and 3

In comparing the calculated value of F_c with the value of F in the table, it is found that the additional term is not significant. This means that, according to this test, the quadratic equation does not result in a better representation of the response surface than the linear equation. But this is a contradiction of earlier findings. Which of the statistical tests can be believed? What can be concluded? If statistics is a precise science, how can it result in contradictory conclusions? Can we not have confidence in statistics? Some of these questions are answered in the following section.

A Priori and A Posteriori Information

Statistics is not a science in itself. Rather it is a tool to aid in scientific studies. Isolated from reality, statistics produces little information. For this reason, it is necessary that researchers know their data and have a sound base in the theory of their science. Above all, experience is highly valuable to the researcher when analyzing and interpreting information collected from experiments and other research.

In Figure VI-8 the rice data and the two curves which were calculated are plotted. The problem is to choose between the two curves -- is the quadratic or is the linear more representative of reality? Table VI-4 is a summary of the statistical values calculated for the two equations. The value of F_c for each of the equations is similar and both are significant at the 5% level of confidence. However, the F test of the quadratic term indicates that it does not significantly improve the representation of the response phenomenon. This is supported by the lower value of t of the quadratic coefficient, which is significant only at the 10% level. On the other hand, the variance or standard error of the

regression for the quadratic equation ($S_{y.12}2$) is half that of the linear equation ($S_{y.x}2$) and the coefficient of determination R^2 for the quadratic is much better than the R^2 for the linear equation.

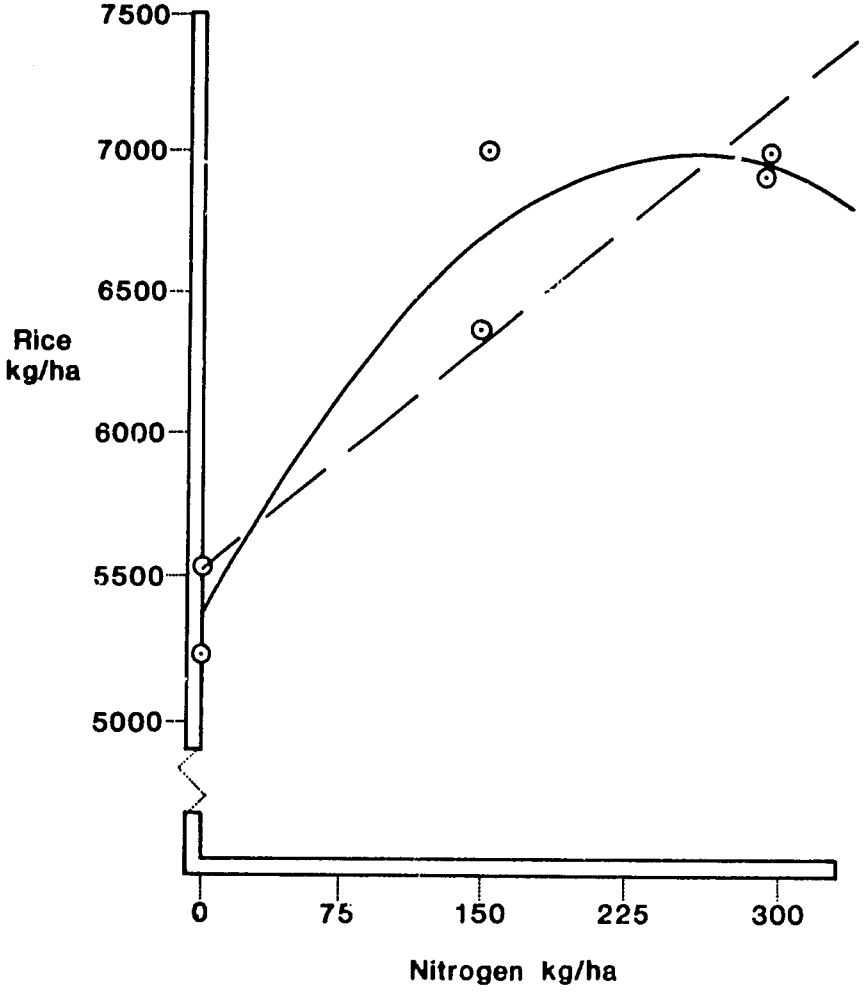


FIG.VI-8. Comparison of linear and quadratic response surfaces.

TABLE VI-4. Summary of statistical values calculated for the linear and the quadratic equations.

Statistic	Linear	Quadratic
R^2	0.813	0.920
$S_{y.x}^2$ or $S_{y.12}^2$	0.152	0.0868
t_b	4.1616**	----
t_{b_1}	-	3.4590**
t_{b_2}	-	-2.0160*
Fc (equation)	17.38**	17.21**
Fc (quadratic term)	-	4.0058 N.S.

**Significant at a 5% level
* Significant at a 10% level

On the basis of pure statistical analysis, the quadratic surface might be rejected in favor of the linear surface even though the R^2 is better for the quadratic surface. The rejection of the quadratic surface would be based on the t test of the coefficient b_2 and of the F test of the same term. The t test indicates that the coefficient b_2 is significantly different from zero only at the 10% level, and the F test indicates that including the quadratic term does not improve the estimate of the response.

Theory and experience with surfaces, and in particular, response to nitrogen, assure us that over a wide range of application, response should not be linear. In this case, the range of application of nitrogen appears to be adequate to give a curvilinear response, and it does appear that the mean values of the observations fall on a curve, not on a straight line. On the other hand, it is possible that experimental results indicate a curve that is more complex than a quadratic curve. That is, within the range of application in the experiment it is possible that all three stages of production exist. However, with only three treatment levels in this experiment it is not

possible to make this kind of conclusion.

Then what can be concluded? Statistical calculations have resulted in a curve, but with tests of significance which lead one to reject the curve. Even though the statistical tests are not significant at high levels such as 1%, 5%, or even 10%, if one were to conclude that the curve is more significant than the straight line, it would be a correct conclusion in more than 80% of the cases. In this specific case, the calculated value of F_c indicates that the quadratic term is significant at a level between 15% and 20%.

To make a conclusion based on the data presented and with theoretical logic and experience, the best evidence available indicates that the response of IR-22 rice to nitrogen with 40 kg/ha of P_2O_5 and 40 kg/ha of K_2O in Corinto, Cauca, is curvilinear between 0 and 300 kg/ha of N. Furthermore, it appears that production reaches a maximum near 270 kg/ha of N. Using a curvilinear equation it is possible to calculate the profitability of capital invested in fertilizer and to determine optimum doses for different combinations of prices for fertilizer and rice. If, on the other hand, one concludes that the linear response is the correct response, these kinds of interpretations cannot be made. The only conclusion that can be reached from the linear equation is that it pays to apply 100 or 150 kg/ha of N, but it also pays to apply 200 and 300 kg and even more. This type of conclusion is not very reasonable.

It can be seen that theory, logic, experience, and other a priori and a posteriori information provide much of the basis for analysis and interpretation of data, and that statistical analysis is not sufficient by itself. In summary, in addition to statistical analysis, the researcher must count on: 1) a priori information such as theory and experience; and 2) a posteriori information such as the physical and economic feasibility of the results.

MULTIPLE REGRESSION--VISIOGRAPHIC

A visiographic method can also be used to estimate a response surface from a two-input experiment, if several levels of each input are included. However, the method is

graphic and does not result in a mathematical equation. If the data are reasonably good, the method can be precise, and with practice different persons will achieve similar results. The procedure is amenable both to agronomic and economic analyses. Data from a rice trial in La Virginia, Riseralda, Colombia, provide an example.

Table VI-5. Partial results from an N-P-K experiment with IR-22 rice in La Virginia, Risaralda, Colombia, conducted by The National Rice Program, Colombian Agricultural Institute (ICA).

Treatments			Replications		
N	P	K	I	II	Mean
			kg/ha		
75	20	20,60	5435	5170	5302
75	60	20,60	6275	4535	5405
225	20	20,60	5715	6330	6022
225	60	20,60	6805	6130	6468
150	40	0,40,80	6540	5907	6223
0	40	40	4440	4500	4470
300	40	40	6950	6800	6875
150	0	40	5400	6450	5925
150	80	40	5670	5280	5475

Source: Hildebrand (1972)

The experimental design of the original trial was a partial or incomplete factorial arrangement in a modified central composite design (see Chapter V) with five levels each of N, P, and K. Response to N and P were stronger than response to K, so for purposes of this example differences in K levels will be ignored. Partial data are shown in Table VI-5.

The first step is to construct a graph with the axes representing quantities of P_2O_5 and quantities of N (Fig. VI-9). The mean yield levels for each known combination of N and P are then plotted on this graph. These values represent different levels on the response surface. Then, with these values as guides, topographic contour lines are

drawn visiographically (Fig. VI-9). Unfortunately, a zero-N, zero-P treatment was not included in this trial, so no value is available for the origin. However, with care, reasonably smooth contours can be drawn. In this example, contour levels of 5000, 5500, 6000, 6200, 6400, 6600, 6800, and 7000 were drawn. It appears that the maximum yield level will approximate 7000 kg/ha.

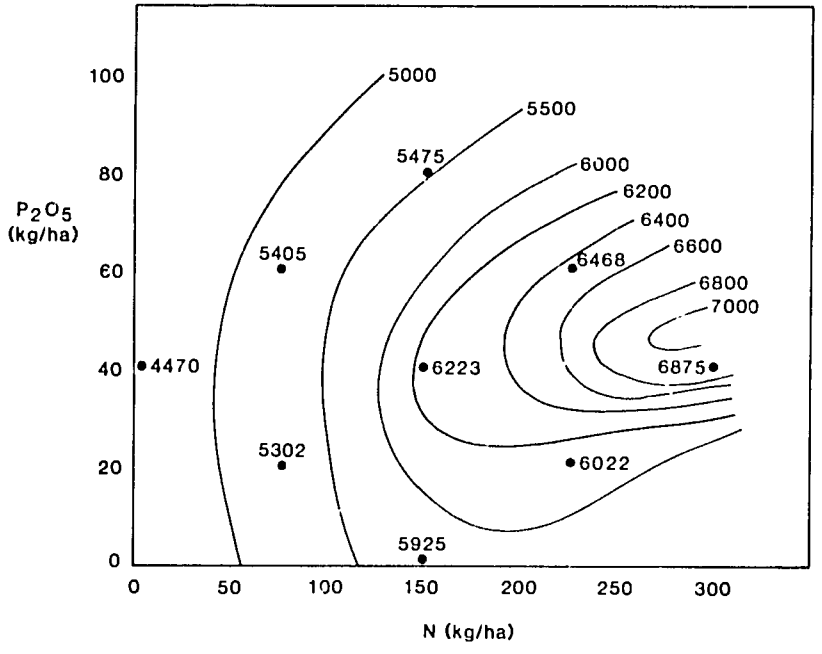


FIG.VI-9. Response surface estimated visually.

The contours in Fig. VI-9 can be smoothed by another simple process (Fig. VI-10). Slices of the response surface in Fig. VI-9 are taken for several levels of P. These represent single variable response surfaces for N. The points estimated from the surface are plotted as in Fig. VI-10, and then a smooth curve is drawn through the

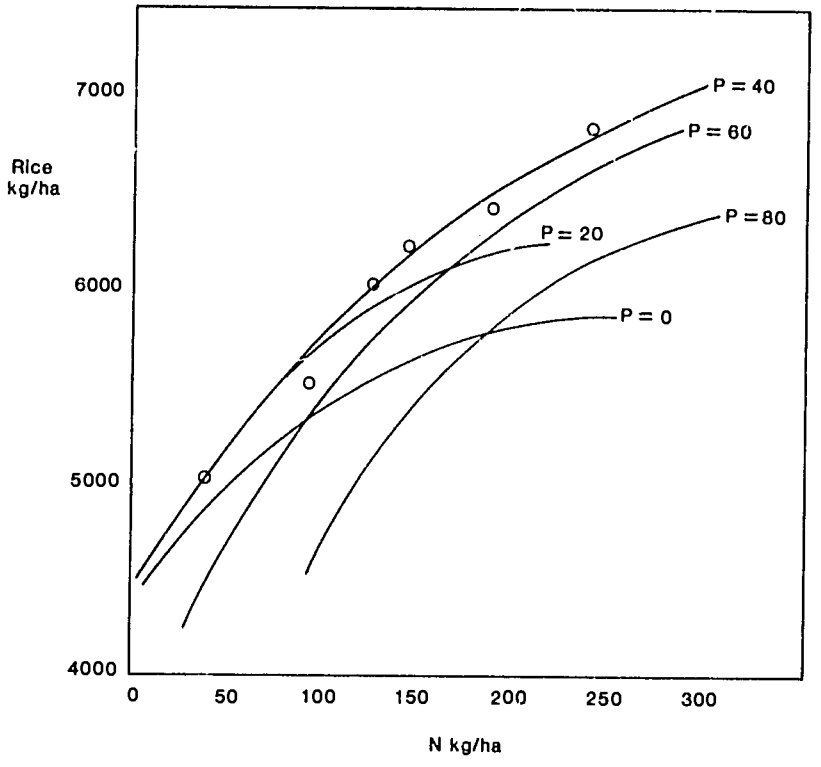


FIG.VI-10. Smoothing the surface with levels of phosphorus.

points, as illustrated for P = 40 in Fig. VI-10. The contours in Fig. VI-9 are then smoothed by taking the N and P values from Fig. VI-10 for each contour level. Fig. VI-11 shows the smoothed contours.

The graphic response surface in Fig. VI-11 can be used for economic interpretation. Consider the price of P as $2.00 and the price of N as $4.00. Budget lines can be constructed as in Fig. VI-11. The first budget line (the furthest to the left) represents a cost of $200. If a farmer has only $200 to spend for fertilizer, he can purchase any combination of N and P along the $200 budget line. The highest yield which can be achieved with $200 of fertilizer is approximately 4950 kg/ha. This would be

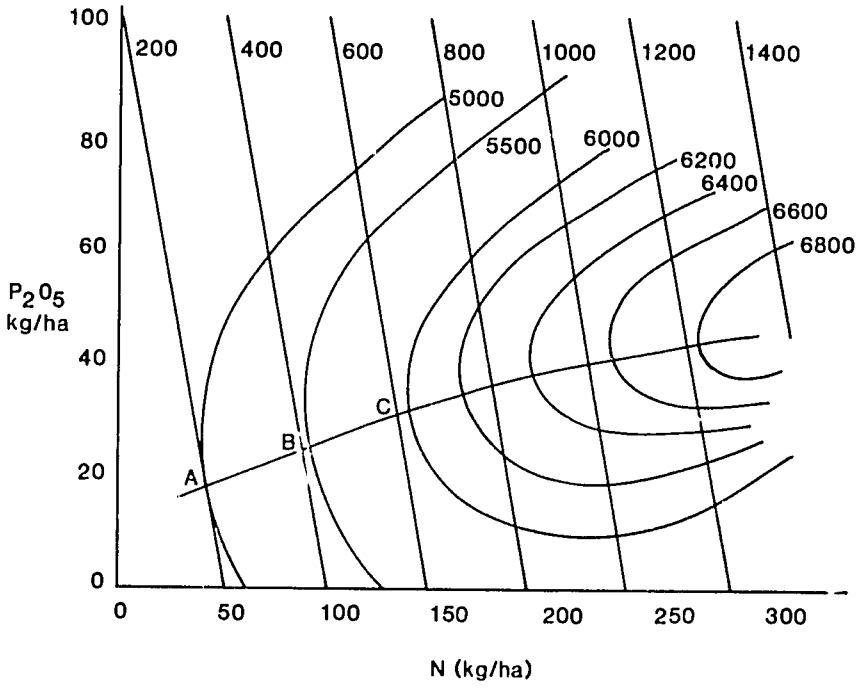


FIG.VI-11. Smoothed product contours and budget lines.

achieved with approximately 37 kg/ha of N and 18 kg/ha of P, represented by point A on Fig. VI-11. If a farmer has $400 to spend, maximum yield will approximate 5450 kg/ha and will be achieved with approximately 80 kg/ha of N and 25 of P, at the intersection labeled point B. With $600 to spend, maximum yield would approximate 5930 kg/ha and would be achieved by utilization of approximately 123 kg/ha of N and 30 kg/ha of P (point C on the figure). Similarly, best combinations and yield levels can be estimated for other budget amounts, such as $1000 and $1400. The cost of moving from one budget line to the next and the value of added yield can be calculated to determine profit or net return for different levels of fertilizer use (Table VI-6). In this example, taken from an earlier period when fertilizer prices were low, it pays to apply fertilizer to produce nearly 7000 kg/ha of rice. If the price of fertilizer were double that used in the example, profit would decrease after the $1200 budget line.

TABLE VI-6. Costs and returns from most economic combinations N and P fertilizer for rice, visiographic estimation, 1970.
P_n = \$4.40/kg, P_p= \$2.00/kg, P_r= 2.80/kg.

Levels of N	Levels of P	Levels of Y	Cost of fert.	Value of rice	Net return[1]
kg/ha	kg/ha	kg/ha	Colombian pesos	Colombian pesos	Colombian pesos
0	0	4400	0	12,320	12,320
37	18	4950	200	13,860	13,660
80	25	5450	400	15,260	14,860
123	30	5930	600	16,604	16,004
166	35	6280	800	17,584	16,784
210	39	6550	1000	18,340	17,340
253	43	6775	1200	18,970	17,770
297	47	6900	1400	19,320	17,920

1. Excluding fixed costs; that is, net only of the variable cost of fertilizer.

VII
Farmer-Managed Trials

Previous chapters have described a process by which researchers evaluate technological alternatives over one or more locations and years, utilizing their own understanding of farmers' resources, desires, and capabilities, and with increasing farmer participation in researcher-managed trials. After analyzing results from previous trials, the best alternatives are then placed in farmers' hands for their evaluation. Keeping in mind that it is the farmer who must ultimately make decisions concerning adoption or rejection, farmer-managed trials (FMTs) provide the opportunity for farmers to become the primary evaluators of new technology. In order for the farmer to be able to evaluate the results of a trial, the trial must possess three critical characteristics: 1) the technology must be simple enough for farmers to comprehend and manage it; 2) farmers must use their own resources so they can understand all implications of the alternatives; and 3) design of the trial must be simple enough that farmers can observe differences in treatments and/or measure them, with their own means of measurement. An example of an FMT might be the testing of a new cultivar under the farmers' normal planting and cultivating procedures. The farmers pay all their usual costs plus the cost of the seed of the new variety. Design of the trial would be simple: farmers plant one plot with the

Previous Page Blank

new variety and a similar check plot with their usual variety.

FMTs can be designed to provide agronomic and economic data for the researcher, but the trial process must not interfere with the farmers' normal activities and their ability to interpret results as they see them. Because there is little control of variability, a large number of farms across a region are required for researchers to measure treatment effect. Regional analysis will determine the level of stability of the new technology and the degree of homogeneity of the recommendation domain. It is important to recognize that trials "lost" for purposes of researcher evaluation are not necessarily lost for the purpose of farmer evaluation. Farmers will decide on the merits of the alternatives without drawing on the kind of technical data needed by the researcher.

Conducting an FMT is relatively easy because of its simplicity. It has a reduced number of treatments (sometimes just one), and usually has no replications. The simplicity of FMTs makes it possible for other institutions to play a role in the evaluation process in the region. For example, being involved in on-farm trials provides an opportunity for extension personnel to become familiar with the technology. A major advantage of this involvement is the establishment of stronger institutional linkages, which can accelerate adoption of newly developed technology.

Recommendation domains are tentatively defined in the initial characterization and are redefined or partitioned during the research procedures described in previous chapters. For purposes of the FMTs, only farmers who fall within the specific recommendation domain for which the alternatives were designed should be chosen to participate.

DESIGN

Because there are large numbers of farmers involved in FMTs and because the trials are not managed by the researcher, there is a limit to the number of measurements the researcher, personally, will be able to make during the trials. The most important, and usually the only one, will be yield. Other information will be available from

crop records maintained in collaboration with the farmers.

Check Plots

At each location (farm) there must be a means for comparing the new technology with that of the farmer. This may come from a check plot or from sampling the farmer's yield. Both of these methods have advantages and disadvantages. One of the main advantages of the check plot is that it is identified ahead of time on a representative part of the farm, and it can be harvested at the same time as the treatment plot or plots. When a

PRACTICAL FIELD ADVICE
Farmers may treat identified check plots differently from their own fields for a couple of reasons. 1) They usually <u>know how</u> to produce more than they do produce and want to prove this to the researcher. The reason they do not produce more on their own fields should be investigated. Results will usually reflect constraints that are imposed on a whole farm but that can be relaxed on a single small plot. The tendency to prove something to the researchers is reduced as confidence builds between researchers and farmers. 2) Because check plots are part of the "trial," farmers may wait until researchers tell them something needs to be done before doing it and not do it at the same time they do their own fields. Good communication between researcher and farmer reduces this problem. The first case results in higher yields in the check plot than on the farmers' own fields. The second case results in lower yields on the check plot. Both kinds of error have been observed.

sampling procedure is used, it may slow harvest because sample sites must be harvested prior to the regular harvest of the farmer. If delays are encountered, the

farmer may harvest the whole field before data are collected. An advantage of the sampling procedure is that the farmer has no means of previously identifying the check area, so it cannot be treated differently from the rest of the field.

If check measurements are to be taken by sampling, it is important that the locations be chosen completely at random. One convenient method is to choose a row and a distance into the field (in that row) from a table of random numbers, and to harvest a set distance in that row (probably the net row length in the trial plot). The procedure is repeated until sufficient replications have been made to make the area harvested by sampling equal to the net size of the treatment plot. Another method of measuring the results of the farmer's practice is to locate a check plot on paper, positioned randomly in the field, and mark its location at the time of harvest.

Plots should be large enough to allow realistic estimates of labor requirements and other economic factors. They should also facilitate evaluation by farmers. The unit or units used should be related to those commonly used by farmers in the area (hectare, manzana, cuerda, tarea, etc.). It is important that the area be large enough on each farm that it can be of significance to the farmer. If possible, 10% to 20% of the area allocated for that crop on each farm should be included, so that the farmer <u>must</u> give it appropriate attention.

Number of Sites

The number of sites needed in an FMT to test an alternative in a homogeneous recommendation domain will depend on several factors, including:

1) the nature of the alternative to be tested;
2) the budget available to researchers;
3) demands on their time for other kinds of trials; and
4) the number of farmers in the recommendation domain.

Usually, the fourth factor is not limiting. The first factor is important because a more complex technology or alternative requires more learning time from each farmer, and the researcher will have to visit each farm more often for supervision.

Field researchers will be involved in several kinds of trials and must regularly spend time at those they are managing. They need time to analyze data and write reports. They require transportation to be able to move among widely dispersed trial sites. All these factors modify the number of farmers each researcher can include in FMTs. For a team of three to five researchers operating in one recommendation domain a minimum number of farms participating in on-farm trials is probably around 20.

ANALYSIS AND INTERPRETATION OF RESULTS

There are four different kinds of evaluation that can be made from farmer-managed trials:

1) the farmers' personal evaluations, from which they decide whether to accept the technology, reject it, or continue testing it;
2) a survey evaluation by researchers in the year of the trial to ascertain what farmers say about the alternative tested;
3) a technical evaluation made during the year of the trial from data collected by researchers and/or farmers; and
4) an evaluation of acceptability made the year following the trial, based on farmers' active acceptance or rejection of the alternative.

Farmers' Evaluation

This is the most important evaluation in an FMT. Farmers know their land, climate, and families' needs and capabilities and have a built-in method of evaluating new technology based on their interpretation of what they see or measure in the trial (in some cases taking into account opinions from neighbors). From the time they need to obtain any new inputs (e.g., fertilizer, seed) or follow a new procedure, their evaluation mechanism is in operation. It is not always necessary for farmers to measure the harvest from a plot in order to make a decision about the acceptability of an alternative. When farmers do measure the harvest, they will be more interested in their own means of measuring than those of the researcher. Researchers' needs should not interfere with farmers' evaluation.

There are two phases in farmers' evaluation. One, passive evaluation, is accomplished the year of the trial. Farmers decide what they think they will do with the new technology the next year (accept, reject, or continue experimenting with it). The second phase is active evaluation. This occurs the year after the trial, when the farmers actively decide whether or not to use the new technology, of their own volition, with all the associated costs and risks.

Passive evaluation

It is often convenient or necessary for an FSR/E team to make a formal evaluation of how farmers feel about the alternatives they are testing during the year in which the test was made. The purpose is to measure the potential acceptability or rejection of new technology to determine whether it may have to undergo further evaluation in FMTs the following cycle or is ready for wider distribution. If it is ready, it will be necessary to assure the availability in the market of any components to be purchased; or assure that there is a market outlet for the product. If reaction from farmers is negative, pertinent feedback to the research team is necessary for modification of the technology or to recommend its elimination from further testing. Farmers' passive evaluation of acceptability can be assessed by a simple directed survey of some or all of the participating farmers to ascertain what they think about the alternatives. The farmers might indicate which of two tested varieties they consider more resistant to a certain disease, which resulted in the best yield, which had the best eating quality or was better for storage, etc. Most important, perhaps, is whether the farmers think they will implement the new alternative the following year. It is also important to ask farmers interviewed why they think they will or will not use the alternative the next year.

In the eastern part of Guatemala, ICTA developed an early, high-yielding, open-pollinated white maize variety that had performed well in on-farm tests for three years (Table VII-1). The new material, B-5, was compared with H-3, a popular, high-yielding hybrid from El Salvador, and with Arriquin, a local early variety.

TABLE VII-1. Maize yields from on-farm trials in Jutiapa, Guatemala.

Genotype	1978	1979	1980	Avg.
	tons/ha			
B-5	4.56	2.80	3.72	3.69
H-3	3.68	2.61	3.69	3.33
Arriquin	3.15	2.48	3.17	2.93

Source: ICTA

In 1981, there were 25 farmers collaborating in this FMT (Samayoa, 1980). A preliminary survey to ascertain farmers' passive evaluation was conducted before harvest. Half the collaborators had planted second or third generation H-3 and other hybrids from seed they had saved. Six planted Arriquin, and the remaining collaborators planted other local varieties. Results indicated that 71% of the farmers said they intended to plant B-5 the following year. More than one-third indicated that their neighbors wanted B-5 seed. Those collaborators who did not indicate that they would plant B-5 seed were waiting for the harvest before deciding. (It can be seen that it would have been better to conduct the survey after harvest but other demands on the researchers at that time prevented it.) In other results, 74% considered B-5 to be resistant to drouth, 63% thought the height of the plant was adequate (not too low, in this case), 75% liked the size of the ear, and 83% agreed that the material was earlier than the variety they planted. As a result of this evaluation, it was recommended that B-5 seed be multiplied so that sufficient amounts would be available the following year for interested farmers.

Active evaluation

In the year that farmers are participating in an FMT, they are continuously making judgements about the technologies being tested. When questioned about their evaluations, they will respond according to how they are thinking at that particular moment. But the real test of the technology is its acceptability. This comes at the time when farmers actually decide whether or not to use the alternative. Even though they may previously have

indicated their intent to use it, now that they must go to town to buy new seed or a new product, or must spend more time working in the field in order to use it, they may decide against it. Or, they may decide to use the alternative but to modify it; for example, they may use only half as much fertilizer as recommended and tested the previous year.

Through another directed survey the year after the FMT, researchers can obtain the information necessary to make an assessment of farmers' active evaluation of acceptability. Farmers are asked if they are using the alternative and, if so, on what proportion of the area of that particular crop for which it was recommended. An index of acceptability can be calculated from this information:

$$I_a = (C \times A) / 100$$

where: I_a = Index of acceptability

C = The percentage of the farmers interviewed who used the practice on at least part of the crop the year following the FMT in which they participated

A = From among those farmers who used the practice the next year, the percentage of the area they have planted to that crop on which they are using the practice

As an example, of 60 farmers who collaborated in an FMT the year before 40 were interviewed for purposes of calculating an index of acceptability. Of the 40, there were 20 (C = 50) who used the alternative on 70% of their crop (A = 70). Thus,

$$I_a = (50 \times 70) / 100$$
$$= 35$$

Experience has suggested that if I_a exceeds 25 and C is equal to or greater than 50, there is a good possibility that adoption of the technology in the recommendation domain will follow. Note that this is an evaluation of acceptability of technology to the farmers.

It is not a measure of "acceptance," which would imply adoption or impact. However, a high I_a can be used to predict a high rate of adoption.

Results of a directed survey to evaluate active acceptability can provide additional information. For example, consider the following possibilities:

	C	A	I_a
Case I	90	10	9
Case II	10	90	9

In Case I, farmers were generally interested in the alternative (90% continued using it) but they were not completely convinced and continued experimenting with it (they used it on only 10% of their crop). In this case, it would be worthwhile for the researchers to continue working with the farmers for another year on the same alternative and to make follow-up evaluations the second year. In Case II, 90% of the farmers rejected the alternative, so it cannot be considered "acceptable" as it is. However, there is a subgroup who considered the alternative very acceptable because they planted most of their crop with the new technology. Therefore, the researchers should define the characteristics that differentiated these farmers from those who rejected the technology and consider partitioning their recommendation domain with respect to this alternative.

A second use of the survey the year after the FMT is to ascertain <u>why</u> farmers accepted or rejected the alternative or alternatives. This information can be used to help guide the development of other alternatives and usually provides additional insight into problems and concerns of target farmers. In the case of a rejected alternative that shows great potential to increase the productivity of a farming system, a further analysis of why it is not acceptable will guide researchers both on- and off-station in how to improve it to meet the farmers' demands. This may be the case of a maize variety for which the ear is not completely covered, causing a high incidence of ear rot. The course of action here would be a breeding project to remove that defect.

In September and October 1980, a sample of 29 collaborators from a group of 35 involved in an FMT for new upland rice varieties was surveyed in eastern Guatemala (Ruano, 1981). The farmers had planted the new varieties the year before and compared them with their own varieties, using their own cultivation and management practices. Of the 29 farmers surveyed, 27 (93%) used the new varieties the following year. These 27 farmers planted 81% of their rice in 1980 with the new varieties. The index of acceptability is calculated as follows:

$$I_a = (93 \times 81) / 100$$
$$= 75$$

This high value indicates that the new varieties possess characteristics that are quite acceptable to the farmers in the recommendation domain. (Unfortunately, this report does not differentiate between the two improved lines tested in the trial.) Following are the major characteristics and reasons given by the farmers for using or not using the new varieties the second year.

Reasons for using	No. of responses (n = 29)
Excellent yield	13
Easy to thresh	10
Appropriate growing period	5
Others	10
Reasons for not using	
Could not find seed	3
Shatters easily in the field	1
Less straw than local varieties	1

The reasons given by the farmers for using or not using the new varieties support the high value of the acceptability index. These varieties can be promoted on a large scale with confidence that they will find a high degree of acceptability and ultimately be widely adopted in the recommendation domain.

In another area, 30 farmers participated in an FMT in maize in 1979, and 27 were interviewed in 1980 to ascertain their evaluation of varieties and fertilizer practices (Ruano, 1981). The results for the fertilizer practices are in Table VII-2. The recommendation for

fertilizer was to apply 97 kg/ha (150 lbs per manzana) of 20-20-0 at planting and 45 kg/ha of urea 25-30 days after planting.

TABLE VII-2. Index of acceptability for maize fertilization practices.

	Applications 1st	2nd
% of collaborators that fertilized as recommended	41	41
% of their maize area fertilized as recommended	32	47
Index of acceptability	13	19

Of the 27 farmers interviewed, six of them (22%) did not use fertilizer at all in 1980 because the price was too high and/or they did not have the money for it. Another 10 farmers (37%) applied some fertilizer, but not at the recommended rates. The remaining 11 farmers (41%) made both applications. They used recommended rates on 32% of their maize in the first application and applied recommended rates to 47% of their maize the second time.

The recommended practice for fertilizer application does not appear to be as generally acceptable as were the rice varieties in the previous example. Perhaps there is a difference in economic capacity or there may be a soil effect that accounts for the difference found in acceptability. There may be two recommendation domains involved so that further analysis of the situation appears warranted. A directed survey or an analysis of field notes and field histories may be used to determine whether that is the case and to help partition the recommendation domain.

Researchers' Evaluation

Because only the treatment variables are controlled, technical analysis of the results of FMTs requires different procedures from those used for the usual experimental designs. A useful method of analysis is to combine a modified stability analysis and a frequency

distribution of confidence intervals. Modified stability analysis is based on a procedure long used by plant breeders. It is a means of using the variability found on different farms to help confirm recommendation domains or indicate a need to partition the domain. The modified stability analysis is a relatively simple procedure which utilizes an environmental index created by on-farm trial data as a means of measuring all the factors which influence response to a technology. These factors include climate, soils, and farmer management practices. A simple associated procedure, a graphic distribution of confidence intervals, is used to evaluate the variability in results to be expected from a technology within a recommendation domain. Both procedures are amenable to calculation with simple calculators; neither requires, though both can benefit from, replicated designs. They are specifically useful for on-farm trials conducted at several locations.

Modified stability analysis

To understand modified stability analysis (Hildebrand, 1984), consider farmer-managed trials conducted over a large number of farms within one preliminary recommendation domain and utilizing an improved cultivar compared with a local variety. No other changes are made from farmers' usual practices. The only constant at each location (farm) is the cultivars. Each farmer subjects them to different soil conditions, planting dates, pest control, fertilizer, and other management practices. A farm for which the average yields of the two cultivars is high (for whatever reason) is considered to be a "good" environment for the crop as measured by the average yield. A farm for which yields are low (for whatever reason) is considered to be a "poor" environment. Environment, then, becomes a continuous, quantifiable variable whose range is the range of average yields from the trial. Yield for each of the varieties can be related to environment by simple linear regression, based on the following equation:

$$\hat{Y}_i = a + be$$

where: $\hat{Y}_i$ = yield of variety i, and

e = environmental index equal to the average yield of all treatments at each location

By fitting the equation independently for each variety, then plotting the yield response to environment for each variety on the same graph, it is possible to visually compare varieties. Using the same procedure, it is easy to generalize these equation sets to any number and kind of treatments.

TABLE VII-3. Maize yield from farmer-managed, on-farm trials, Phalombe, Malawi, 1981/1982.

Trts	Farmers in first village								
	1	2	3	4	5	6	7	8	TMV
	t/ha								
LM	2.2	2.2	1.9	1.2	1.3	0.9	1.0	0.5	1.4
LM-F	3.6	3.7	4.3	3.2	2.3	2.3	3.1	2.8	3.2
CCA	3.5	2.0	2.9	0.4	0.6	0.5	0.6	0.3	1.3
CCA-F	5.0	4.7	4.3	3.5	2.4	1.7	3.0	2.8	3.4
MF = e_i	3.6	3.2	3.3	2.1	1.7	1.3	1.9	1.6	

	Farmers in second village						
	1	2	3	4	5	6	
LM	1.8	1.1	1.6	1.0	1.6	0.6	1.3
LM-F	3.2	2.5	2.9	1.2	1.9	0.8	2.1
CCA	2.2	0.7	0.9	0.3	1.1	0.3	0.9
CCA-F	2.9	2.5	2.1	1.1	0.8	0.4	1.6
MF = e_i	2.5	1.7	1.9	0.9	1.4	0.5	

Source: Hansen et al., 1982.

LM = Local Maize
LM-F= Fertilizer Local
CCA = CCA Maize
CCA-F = Fertilizer CCA
MF = Mean for farmer
TMV = Treatment Mean for Village

Data for a maize trial from the Phalombe Project in Malawi (Hansen et al., 1982) are analyzed as an example. Fourteen farmers from two villages participated in trials conducted on their respective farms. A simple, non-replicated 2 x 2 factorial arrangement with two maize

varieties and two levels of fertilizer (0 and 30 kg N/ha) was used. All maize was intercropped with cowpeas and sunflowers, a common practice in the area. Maize varieties tested were the "local" flint variety and CCA, an improved semi-flint composite. The data for the Phalombe Project area are in Table VII-3.

Data for each fertilizer level and for each variety are fit by simple linear regression. This can be accomplished easily on simple, pre-programmed electronic calculators.

The calculations for simple linear regression for non-programmed calculators follow, based on using local maize without fertilizer (LM) as an example. The values of Y_i are the individual farm yields for LM, or 2.2, 2.2, 1.9,..., 0.6. The values of $\underline{e}$ in the equation are the mean yields for each farm, or 3.6, 3.2, 3.3, ..., 0.5. By convention, X is used for $\underline{e}$ in the following.

$$\begin{aligned}\Sigma x^2 &= \Sigma X^2 - [(\Sigma X)^2 / n] \\ &= (3.6)^2 + (3.2)^2 + \ldots + (0.5)^2 - [(27.6)^2 / 14] \\ &= 65.02 - 54.41 \\ &= 10.61\end{aligned}$$

$$\begin{aligned}\Sigma xy &= \Sigma XY - (\Sigma X \Sigma Y / n) \\ &= (3.6 \times 2.2) + (3.2 \times 2.2) + \ldots + (0.5 \times 0.6) \\ &\qquad - [(27.6)(18.9) / 14] \\ &= 42.68 - 37.26 \\ &= 5.42\end{aligned}$$

$$\begin{aligned}b &= \Sigma xy / \Sigma x^2 \\ &= 5.42 / 10.61 \\ &= 0.511\end{aligned}$$

and

$$\begin{aligned}a &= \bar{Y} - b\bar{X} \\ &= 1.35 - 0.511(1.97) \\ &= 0.34\end{aligned}$$

To estimate R^2 it is necessary to calculate Σy^2:

$$\begin{aligned}\Sigma y^2 &= \Sigma Y^2 - [(\Sigma Y)^2 / n] \\ &= (2.2)^2 + (2.2)^2 + \ldots + (0.6)^2 - [(18.9) / 14] \\ &= 29.41 - 25.52 \\ &= 3.89\end{aligned}$$

and

$$R^2 = b\Sigma xy / \Sigma y^2$$
$$= (0.511)(5.42) / 3.89$$
$$= 0.71$$

This equation, and others calculated the same way, along with the data points, can be displayed graphically for visual comparisons (Figs. VII-1 and VII-2). In each case, the R^2 value indicates a very good fit. It appears that the materials respond differently to environment and that the local material is superior in "poor" maize environments while the improved material is superior in "good" maize environments. Both cultivars, however, respond favorably to fertilizer in both good and poor environments.

The analysis provides evidence that the farms are part of two recommendation domains. This evidence warrants further study. In the poorer environments (e < 2), i.e.,

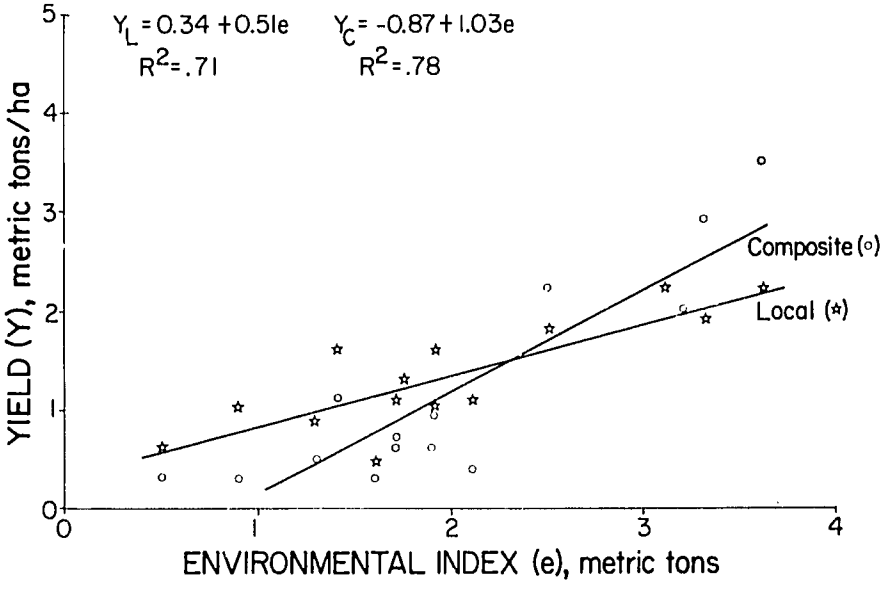


FIG.VII-1. Grain yield response for local maize (L) and CCA composite (C) to environment, without fertilizer, Phalombe Project, Malawi.

those farms which normally do not produce more than 1.5 tons/ha of local maize without fertilizer (the traditional technology), the local material is superior whether or not the farmer fertilized at the rate used in the trial. However, if local, unfertilized maize usually yields more than 1.5 tons/ha on a particular field or for a particular farmer (the better maize environments with e > 2), the new material is superior whether or not it is fertilized at the rate used in the trial. These conclusions apply to a field or a farm.

Although results from two or more years could be used in the analysis, use of the environmental index negates many of the problems associated with only one year's data. It measures response to good or poor environments regardless of the reasons that those environments are good or bad. If another year is better or worse for maize, the data points for an individual farm will shift to the right

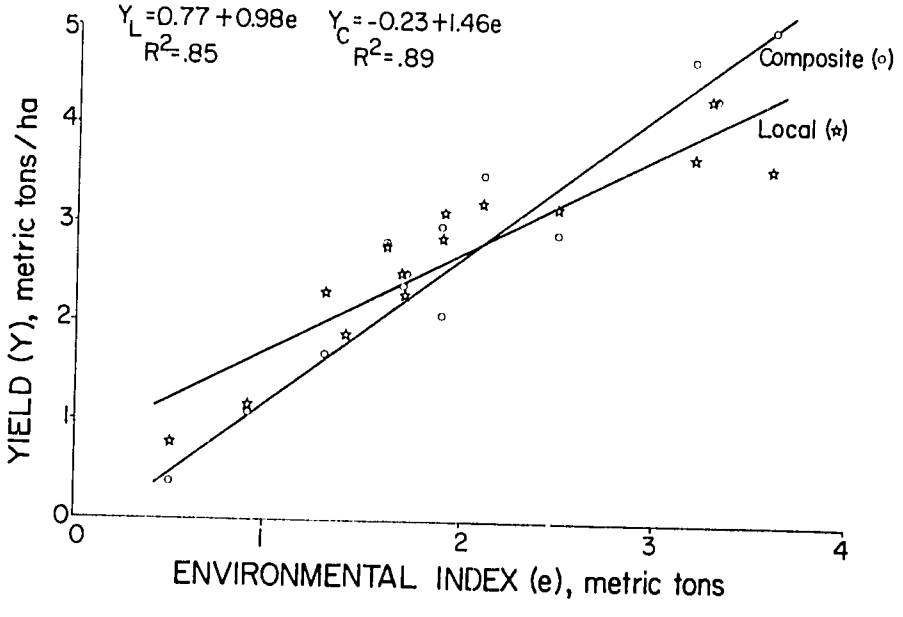


FIG.VII-2. Response of local maize (L) and CCA composite (C) to environment, with fertilizer, Phalombe Project, Malawi.

or left, but a "3" environment changes little if the same treatments are used. Only if the usual environmental index range is much higher or lower, or the range of the index very narrow so that extrapolation is extreme, should there be concern with the use of data from only one year.

Distribution of confidence intervals

It appears from the modified stability analysis that two recommendation domains exist relative to the use of maize cultivars in the farmer-managed trial in Malawi. A method of further examining the data to aid in making specific recommendations is to use a distribution of confidence intervals for the technologies and for each partitioned recommendation domain. In this case there are nine farms in the "poorer" environment ($e < 2$) and five in the "better" environment ($e > 2$).

The confidence interval is calculated for any level of confidence, a, from

$$\bar{Y} \pm t_a S / (n^{1/2})$$

where

$\bar{Y}$ = the mean treatment yield,
a = the level of confidence,
$S = [\Sigma x^2 / (n-1)]^{1/2}$, and
t_a = value from a "t" table

In the example, for $e > 2$ and for the unfertilized local variety, the yields are 2.2, 2.2, 1.9, 1.2, 1.8. As always

$$\Sigma x^2 = \Sigma X^2 - (\Sigma X)^2 / n$$
$$= 17.97 - 17\ 30$$
$$= 0.67$$

$$S = (0.67 / 4)^{1/2}$$
$$= 0.409$$

So the confidence interval for any level a is

$$\bar{Y} \pm t_a(0.409 / 2.236)$$

or

$$1.86 \pm t_a(0.183)$$

From a two-tailed "t" table for n - 1 = 4 degrees of freedom, the following values are obtained.

Degrees of freedom	Probability of a larger value (a)					
	0.50	0.40	0.20	0.10	0.5	0.01
4	.741	.941	1.533	2.132	2.776	4.604

A 50% confidence interval is

$$1.86 \pm 0.741(0.183) = 1.86 \pm 0.14$$

or from 1.72 to 2.00/ha. A 90% interval is

$$1.86 \pm 2.132(0.183) = 1.86 \pm 0.39$$

or from 1.47 to 2.25/ha. This last value means that a farmer with this environment or in this recommendation domain could expect a yield of from 1.47 to 2.25 t/ha 90% of the time from the local variety, unfertilized. By calculating several intervals, smooth curves such as those in Fig. VII-3 can be drawn. In this way, any two (or more) technologies can easily be compared. In Fig. VII-3, when the cultivars are not fertilized, the local variety is much more stable than the composite. Even though there is some potential for higher yield with the composite, the risk of very low yields may be greater than farmers would be willing to accept. If a farmer in the better environment fertilizes, however, then there is a clear choice in favor of the composite. With either cultivar, fertilization presents a superior technology.

In the case of farms in the poorer environment (Fig. VII-4), the local variety is clearly superior, with or without fertilizer. Even here, though, fertilization appears to be a technology that should be recommended if farmers have enough cash to purchase it.

For the two recommendation domains in the Phalombe area, then, the recommendations based on a technical analysis would be to fertilize the local maize variety in the poorer environment and to use the composite maize with fertilizer in the better environment.

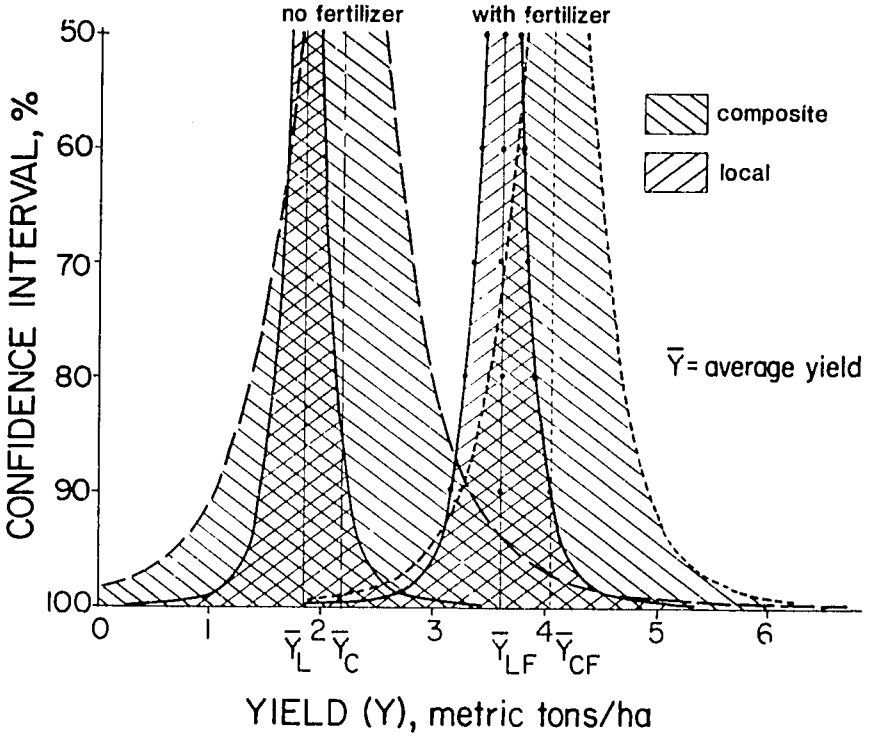


FIG.VII-3. Distribution of confidence intervals for grain yield of local and CCA composite maize, Phalombe Project, Malawi. High environments-five farms where average yield (Y) greater than 2 t/ha.

To be effective for extension purposes, a way to differentiate between recommendation domains must be available. In this case, a study of field histories showed that the better environments were related to the use of manure. Manure, in turn, was available only to the farmers who had cattle. Hence, the better environments, which comprise one recommendation domain, included those farmers who had cattle and who used the manure on a field where maize would be planted.

Another example of the analysis of farmer-managed trials comes from ICTA, in Guatemala. In 1976, following a Sondeo which characterized the western Chimaltenango area and defined three recommendation domains (Duarte et

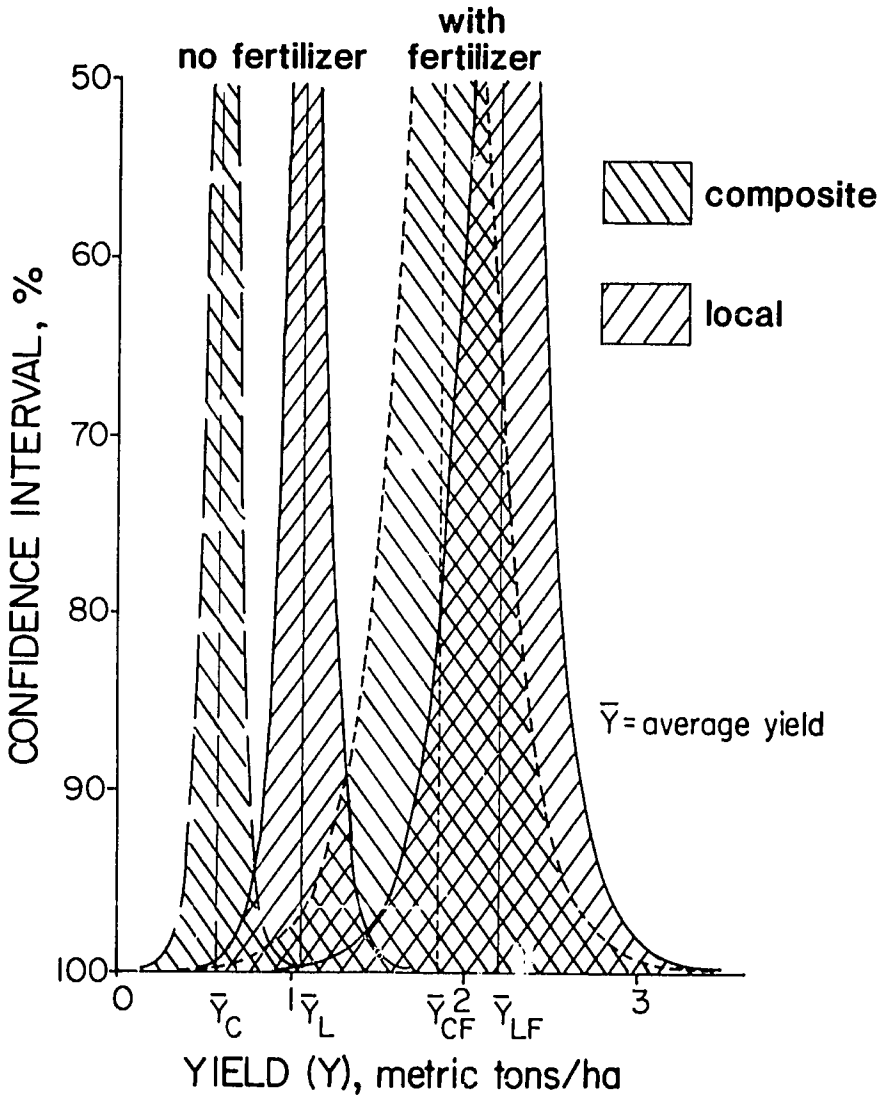


FIG.VII-4. Distribution of confidence intervals for grain yield of local and CCA composite maize, Phalombe Project, Malawi. Low environment-nine farms where average yield (Y) less than 2 t/ha.

al., 1977), a technology was designed that utilized twin maize rows (Fig. VII-5). This technology was designed to maintain the maize yield of the subsistence farmers in the area, who barely had enough land to provide their families' maize needs for the year. Pairs of maize rows were compressed and a bed one meter wide became available between each pair of rows for the production of other crops, which were envisioned to be cash crops (Hildebrand et al., 1977). Maize yields in early trials were comparable to the traditional maize yields and did not require any changes in technology within the maize rows, other than the practice of pairing. Crops grown between the maize rows included common beans, potatoes, and wheat.

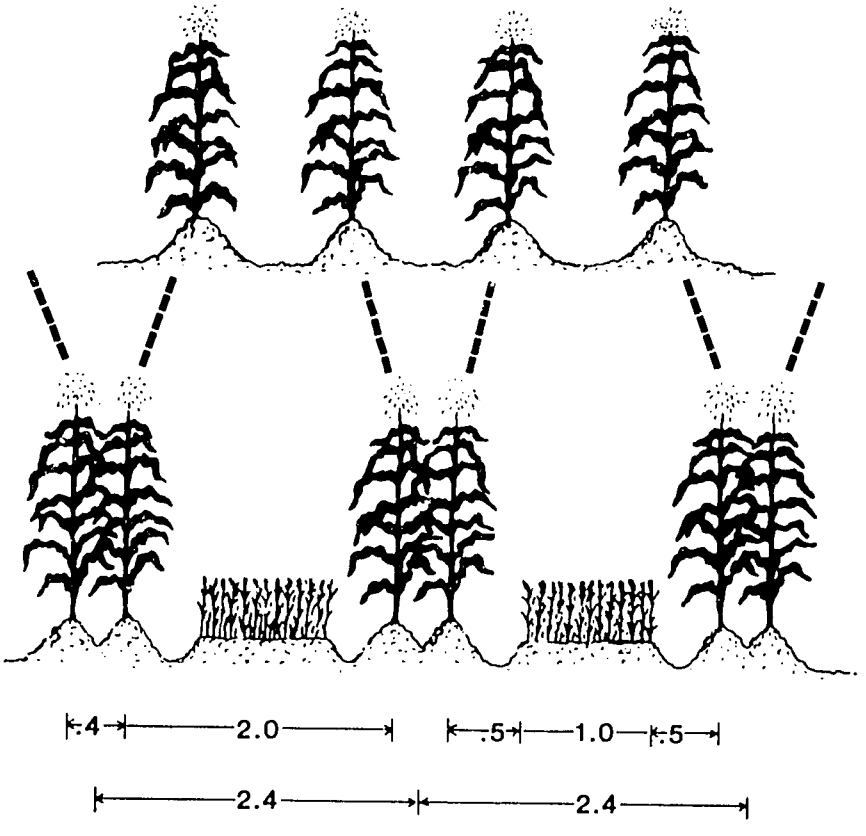


FIG.VII-5. Arrangement of double maize rows. Distance in meters.

In 1979, ten farmers participated in farmer-managed trials utilizing the twin rows of maize and interplanting potatoes. This technology was compared with the farmers' own technology (Table VII-4). Results of the nine trials for which adequate data were obtainable appear in Table VII-5. The results are analyzed with a distribution of confidence intervals for various evaluation criteria.

TABLE VII-4. Technologies compared in farmer-managed maize trials in 1979 in Chimaltenango, Guatemala.

Variables	Farmer's Practices (Avg)	Recommended by ICTA
Variety	Local	Local
Fertilization		
kg/ha N	59	101
P_2O_5	40	64
Plant population (1000/ha)	41.7	41.7
Distance (m)		
Between hills	1.0	1.0
Between rows	1.2	0.4 and 2.0

Source: ICTA, Guatemala

To calculate the confidence intervals, the following procedure is used. Maize yield for the recommended technology (Table VII-5) will be used as an example.

Calculate sum of squares:

$$\Sigma x^2 = \Sigma X^2 - [(\Sigma X)^2 / n]$$
$$= 175.0455 - 169.9547$$
$$= 5.0908$$

mean square: $S^2 = \Sigma x^2 / (n - 1)$
$$= 5.0908 / 8$$
$$= 0.6364$$

$$S_{\bar{x}}^2 = S^2 / n$$
$$= 0.6364 / 9$$
$$= 0.0707$$

TABLE VII-5. Yield, income, and cost from nine locations in a farmer-managed maize and potato trial, Patzun, Guatemala.

Yield		Gross	Total	Net	Net income/
Maize	potato	Income	Cost	Income	Total cost
-metric ton/ha-		———	$/ha	———	Ratio
			Improved technology		
4.66	6.80	2368	1652	716	0.43
4.74	8.29	2710	1495	1215	0.81
3.08	10.34	2851	1482	1369	0.92
3.81	13.76	3740	1497	2243	1.50
4.00	10.75	3113	1493	1620	1.08
5.47	4.86	2092	1496	596	0.40
3.44	5.57	1869	1475	394	0.27
4.98	6.21	2298	1493	805	0.54
4.93	8.10	2704	1495	1209	0.81
			Traditional technology		
3.86		722	343	379	1.10
5.28		987	359	628	1.75
3.48		651	339	312	0.92
3.61		675	341	334	0.98
3.56		666	340	326	0.96
4.41		825	350	475	1.36
3.93		735	344	391	1.14
4.76		890	353	537	1.52
2.65		496	330	166	0.50

Source: ICTA, Guatemala

standard error: $S\bar{x} = (S\bar{x}^2)^{1/2}$
$= 0.266$

or with a calculator that is programmed to provide the standard deviation, S

Standard error: $S\bar{x} = S / (n^{1/2})$
$= 0.7977 / 3$
$= 0.266$

The confidence interval is calculated as follows for various values of a:

$$\bar{x} \pm t_a S_{\bar{x}} \text{ with n-1 degrees of freedom}$$

In the example there are (9-1) = 8 degrees of freedom. For a = 0.50 the value of t is 0.706 so the confidence interval is 4.35 ± (0.706) (0.266) or from 4.16 to 4.54 t/ha.

All confidence intervals are calculated in the same way. Fig. VII-6 shows the distribution of confidence intervals of gross income for the improved system and the traditional system. The traditional system is very stable but at a much lower level of gross income. The same is true of net income (Fig. VII-7). On the basis of these two evaluation criteria, it would seem that the improved system is definitely superior to the traditional system.

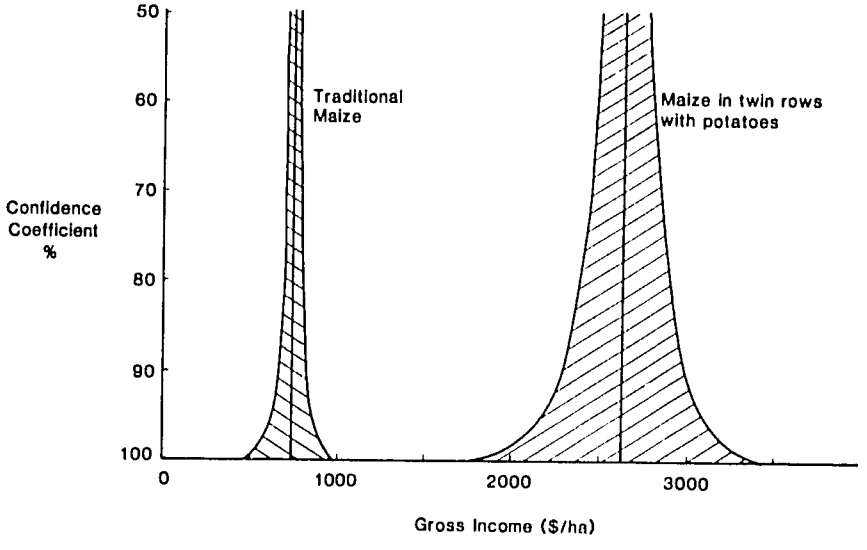


FIG.VII-6. Distribution of confidence intervals for gross income, traditional maize versus maize in twin rows with potatoes.

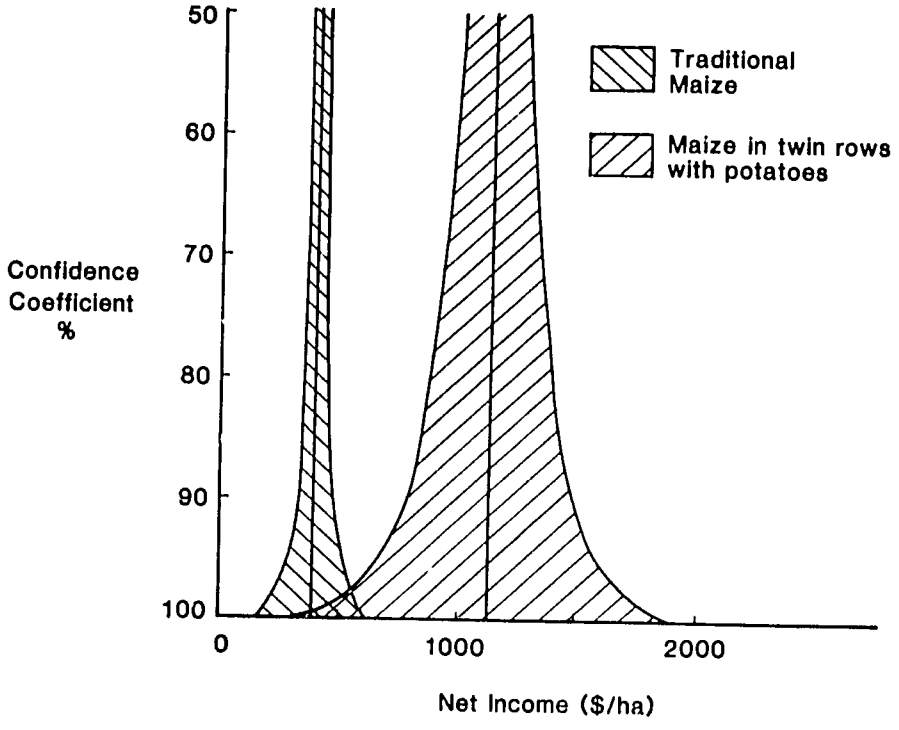


FIG.VII-7. Distribution of confidence intervals for net income, traditional maize versus maize in twin rows with potatoes.

However, there are other criteria which may be much more important for the farmer. In this area, foremost in farmers' minds is producing enough maize to feed their families. For maize yield (Fig. VII-8), which would be important from this point of view, it appears that the improved system does not present any more risk than the traditional system. In fact, it may be slightly superior. So this criterion also appears to be favorable for the improved technology. However, it should be noted that the improved system uses approximately five times as much capital as the traditional system. This is due mostly to the production of potatoes. The criterion of return to investment is completely different (Fig. VII-9). The traditional practice is superior to the alternative being tested. For the farmer who has little capital to invest

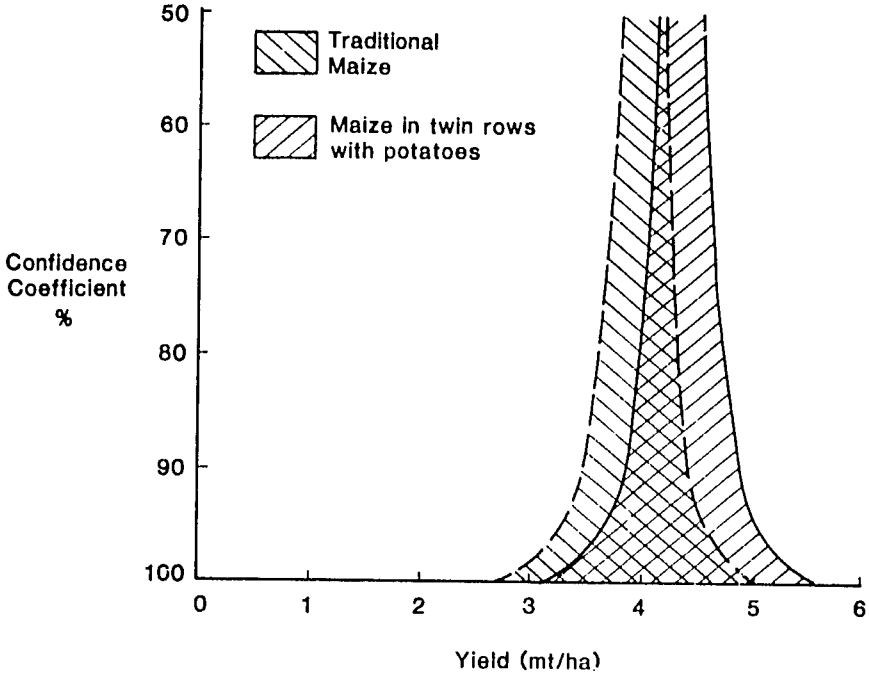


FIG.VII-8. Distribution of confidence intervals for maize yield, traditional maize versus maize in twin rows with potatoes.

and many alternative uses for this capital, there may be sufficient risk to prevent widespread adoption of an otherwise apparently good technology.

The year following the farmer-managed test with this technology, personnel from ICTA returned to interview participating farmers to calculate an index of acceptability for the technology (Table VII-6). Seven of the nine farmers who provided usable data were interviewed at this time. (To have more confidence in the index of acceptability, the number of farmers included should be greater.) Four of the seven (57%) were using the improved system. However, they were using it only on 16% of the land they had planted in maize, so the index of acceptability was only nine. This indicates a fairly low level of acceptability of the practice.

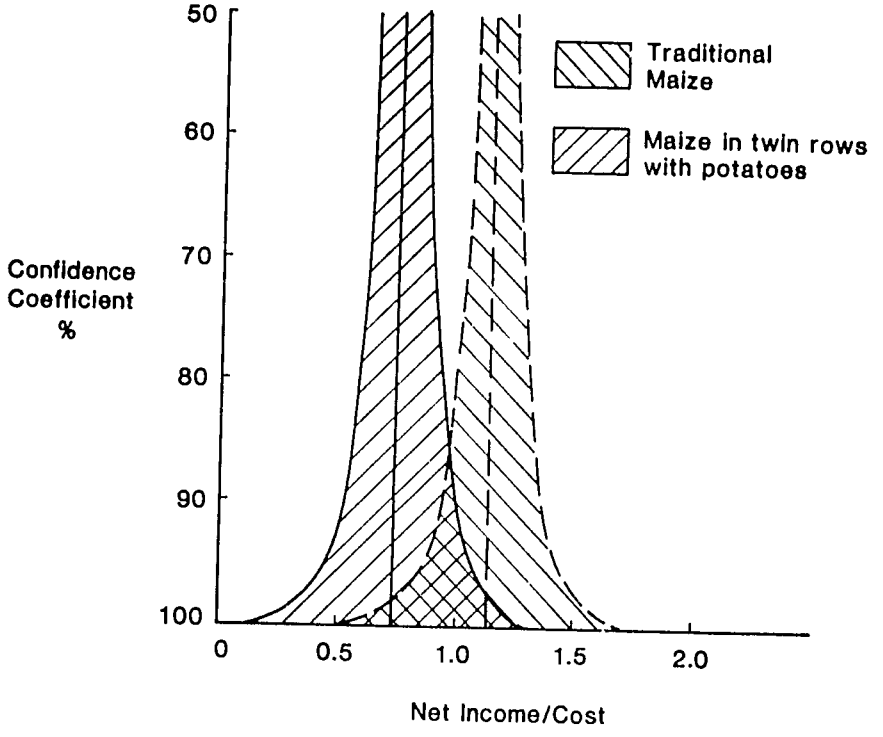


FIG.VII-9. Distribution of confidence intervals for return to cost, traditional maize versus twin rows with potatoes.

If one examines the system involved, it is evident that it is not a substitute for maize, as was implied in the calculation of the acceptability index, but rather a means of allowing farmers to produce an alternative cash crop (potatoes) without reducing the amount of maize produced on the farm. Therefore, a more relevant measure of acceptability would be to estimate the index based on potato production on each farm, rather than on maize production. Data are not available from the directed survey on this criterion. However, it would be anticipated that the farmers who used the system the year following the trial planted a high proportion of their potatoes in the new system. For example, if the only potatoes the four farmers planted were in the new system, the index of acceptability would be 57.

TABLE VII-6. Number of collaborators and area planted in twin rows of maize interplanted with potatoes, Chimaltenango, 1980 (n = 7).

	Area (ha) or Number	Percent
Total area planted to maize (ha)	3.85	100
Collaborators who used the twin rows	4	57
Total area of all maize on 4 farms with twin rows (ha)	1.4	-
Total area of maize in twin rows on 4 farms	0.23	16
Index of acceptability (57 x 16) / 100		9

Source: Samayoa, 1980

Response surfaces derived from modified stability analysis

Previous examples of the modified stability analysis were based on discrete, or non-continuous, variables. Crossing of the regression lines suggests a need to partition the recommendation domain because it indicates that in part of the domain one technology is better, while in another part of the domain a different technology is superior.

Modified stability analysis is also amenable to analysis of continuous variable data, such as fertilizer response. With a continuous variable, differential response between poor and good environments can be detected and used to partition a recommendation domain even where cross-over does not occur. Data from a nitrogen response trial for radishes conducted as part of a farming systems class at the University of Florida are used in the following to show continuous data analysis and the potential for partitioning recommendation domains.

TABLE VII-7. Average radish yields from "farmer-managed" nitrogen response trial, 1984.

"Farm" No.	N Level (cwt/acre) 0	1	2	3	4	$\bar{x}$=e
	tons/acre					
1	1.56	2.00	2.49	1.40	1.35	1.76
2	0.54	1.08	0.91	0.50	0.42	0.69
3	1.68	2.58	1.64	2.17	2.33	2.08
4	0.29	0.43	0.65	0.96	0.74	0.61
5	0.77	1.87	1.65	0.83	1.21	1.27
6	0.92	1.84	1.44	1.04	1.30	1.31
7	1.04	3.72	2.18	2.31	1.95	2.24
8	1.81	2.75	3.00	3.17	2.73	2.69
9	0.98	2.75	2.31	2.83	2.70	2.31
10	0.87	0.83	1.16	0.85	0.90	0.92
11	0.44	1.99	2.24	2.07	1.35	1.62
12	1.01	2.30	1.85	2.21	0.40	1.56
Treatment average	0.99	2.01	1.79	1.70	1.45	1.59

Source: Farming systems class trials, Univ. of Florida

The trial was conducted at two locations on the agronomy farm on the campus in 1984. The class was divided into six teams, each consisting of four members. Each team conducted an experiment using a randomized complete block design at each of two locations. Because the teams managed their trials differently, each trial at each location is considered a farm for purposes of analysis. Hence, there are twelve farms included in the trial. Data are shown in Table VII-7.

The linear equations calculated for each of the treatments are of the form

$$\hat{Y} = a + be$$

where $\hat{Y}$ is yield and e is the environmental index for each farm. The equations as estimated by least squares are summarized in Table VII-8. The equations are shown graphically in Fig. VII-10.

TABLE VII-8. Linear equations for modified stability analysis calculated from radish data.

Treatment cwt N/acre	a	b	R^2
0	0.13	0.54	0.55
1	0.04	1.24	0.81
2	0.31	0.93	0.80
3	-0.23	1.21	0.84
4	-0.26	1.08	0.76

To obtain the data for the response surfaces, estimates of yield for each N level can be taken from the graph, or the linear equations in Table VII-8 can be solved. The latter is more precise. Yields are shown in Table VII-9 for two environmental levels, e = 1.5 and e = 2.5, which lie within the range of data. These values were obtained by solving the equations in Table VII-8 for the two values of e and for each N level.

TABLE VII-9. Estimated radish yield for five levels of N for two environments.

e	cwt N/acre 0	1	2	3	4
	tons/acre				
1.5	0.94	1.90	1.70	1.58	1.36
2.5	1.48	3.14	2.64	2.80	2.44

The data in Table VII-9 are then used to obtain two response surfaces. A quadratic surface seems appropriate, based on the type of response. The two equations, estimated by least squares, follow:

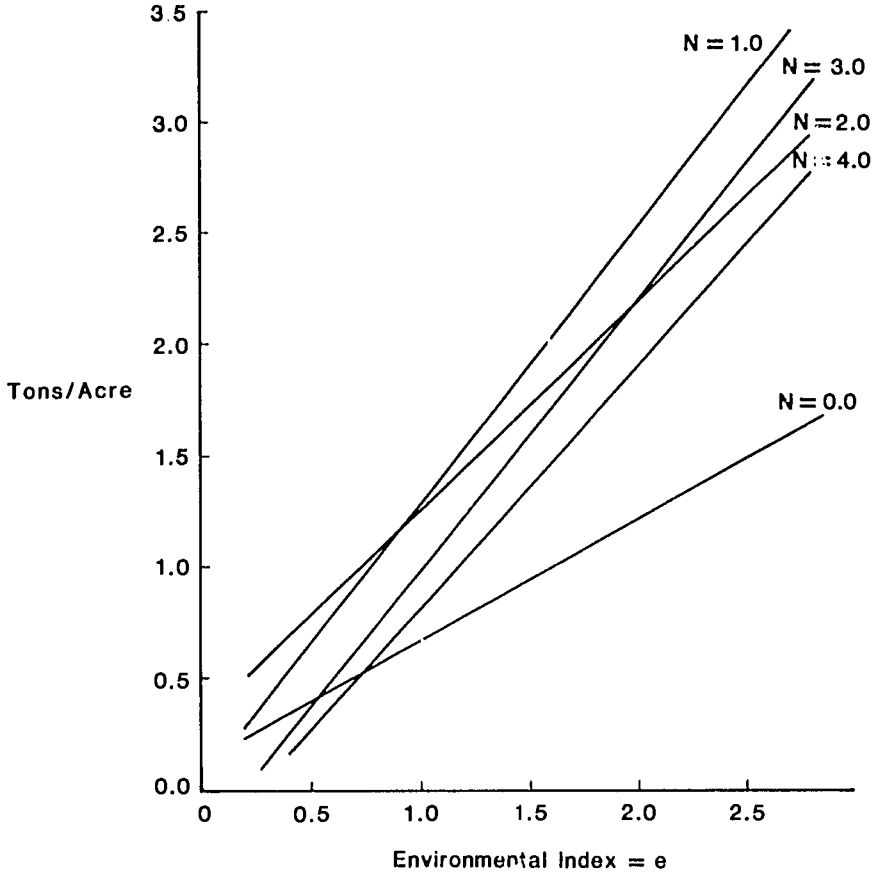


FIG.VII-10. Response to environment of N treatments on radishes.

$$\hat{Y}_{1.5} = 1.07 + 0.72N - 0.17N^2 \qquad R^2 = 0.74$$

$$\hat{Y}_{2.5} = 1.70 + 1.11N - 0.24N^2 \qquad R^2 = 0.68$$

These equations are graphed in Fig. VII-11. Biologically and economically, these two curves appear to represent different situations. In the better environment, production is maximized with 2.31 cwt of N. Maximum yield is 2.98 tons. Maximum production in the

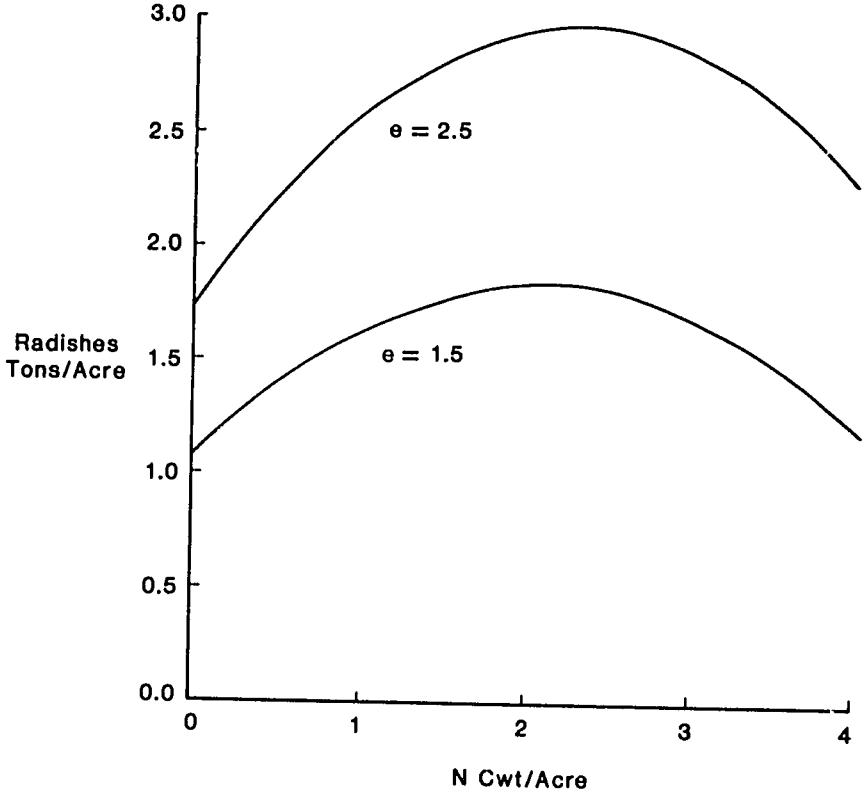


FIG.VII-11. Response of radishes to N for two different environments.

poorer environment, with 2.12 cwt of N, is 1.83 tons of radishes. With prices of $30/cwt for N and $400/ton for radishes, the most profitable application of N is 2.16 cwt in the better environment and 1.90 in the poorer. Yield in each case is 2.96 and 1.82 tons respectively. In the better environment, the 2.16 cwt N cost $64.80 and the increased radish production from adding the 2.16 cwt N is worth $504.00. Profit in the better environment then is $439.20 and net rate of return to investment is 439.20 / 64.80 = 6.78. In the poor environment, the 1.9 cwt of N cost $57.00 and increased production was worth $300.00. Profit in this environment is $243.00 and net rate of return to investment is 243 / 57 = 4.26.

The important aspect of this example is to demonstrate how management, as opposed to soils and climate, affects yield, and how differences in management can be utilized in analyzing results of on-farm trials. Within locations (the first six "farms" in Table VII-7 are from one location and the last six from another), the differences resulted from the way each team planted and cared for the radishes -- soils and climate were essentially identical. Between locations, soils differed somewhat, but climate was similar. For purposes of partitioning a recommendation domain, all of the first location could be considered to be in the poorer environment. But in the other location, half the "farms" are in the better environment and half in the poorer environment. An analysis would have to be made to determine the sources of this difference in order to define the partitioned recommendation domains.

VIII
Initiating and Managing FSR/E Programs

Previous chapters have dealt with technical and statistical aspects of on-farm research in FSR/E. This chapter discusses problems and practices involved in the organization and management of FSR/E programs.

Most countries consider the development of the small farm sector as a high national priority. Existing public and private sector research and farm service organizations often address the needs of the larger, more commercial farms and those in the best agricultural regions. For these reasons, an early and necessary step in initiating on-farm research is likely to be a conscious political and management decision to target a larger share of research and extension resources on the problems and priorities common to small, limited-resource farms. The goals of a national FSR/E system can be broadly stated to be generating, testing, evaluating the acceptability of, and disseminating new or modified technology which can be directly applied by farm families in their farm operations. The adoption of the improved technology would be expected to result in an expansion of production and improved standards of living for farm families.

In describing the organization and management of a national FSR/E system, it is assumed here that a decision has been made by the directors of national research and/or extension organizations to develop an on-farm research

Previous Page Blank

program. The organization and management of FSR/E in a particular country will depend upon the nature and responsibility of existing institutions. In several countries, of which Guatemala is the best documented example, separate institutes were created to carry out on-farm research. In other countries, limited agro-technical and agro-biological research is already being carried out on farms, and some socioeconomic and physical resource information is being collected and analyzed by a wide variety of services, industry groups, and universities. In some country settings, it is the extension organization which is better prepared to conduct on-farm or applied research and which may provide the impetus for an FSR/E program.

GETTING AN ON-FARM RESEARCH PROGRAM STARTED

For most institutions, when adding or expanding an on-farm research program, some important reorganization and reallocation of resources will be required. Most larger research institutions are organized for program implementation through several commodity and disciplinary programs. It is rarely possible financially for each program to initiate and carry out its own on-farm research sequence, although some larger commodity programs may already have a number of varietal trials on farms. From several points of view, it may be preferable to form a separate program group with responsibility for most on-farm research activities. The actual conduct of trials and surveys would be delegated to multi-disciplinary teams assigned to specific regions or geographic areas. Team composition and the process of defining and choosing geographic areas are discussed later in this chapter.

Many research institutions lack social science groups, or, where they do exist, their programs are usually considered to be totally separate from biological programs. Successful on-farm biologic research programs are dependent upon the professional inputs of agricultural economists, sociologists, anthropologists, and extension education specialists. These disciplines are critical but research directors and others in authority may need to assure that they are represented and accepted as full participating members in planning and implementing all programs of work. Each of these disciplines has an

important contribution to make in the selection of research problems and priorities, the design and analysis of results, and the screening of recommendations for farmers. Carrying out multi-disciplinary team work in an interdisciplinary manner is not an easy mode of operation for most persons trained in narrow disciplinary fields. It needs to be encouraged and supported.

Although on-farm research activities are a logical outgrowth of research institutions, it is also possible to have an extension institution involved in applied or adaptive research. This alternative should not be considered as a shift of the research responsibility from a research to an extension institution. The on-farm program should be related to the research institution for technical orientation and support needed by the extension staff.

Close cooperation between research and extension is not common in most developing countries. Often their institutions are budgetarily and physically separated; national extension and research organizations are seldom equal in budget and number of personnel. It is also not uncommon to find a lack of cooperation between these organizations and even personal animosity between researchers and extensionists. All these factors affect the flow of technology transfer from the research program to extension activities. Therefore the alternative to have on-farm research activities within an extension framework is, in principle, one way to establish a rewarding relationship between the two organizations.

PLANNING AND MANAGING ON-FARM RESEARCH

Any successful research program requires some mechanism for establishing a work plan and for periodically reviewing progress. Advanced planning of on-farm research is particularly important because the work occurs off-station. Consequently, it is important to specify, in advance, what research is to be undertaken and the degree of support to be provided by commodity or disciplinary teams, as they must work in a decentralized manner.

The first step in developing a work plan is problem identification. Objectives of the initial characterization and exploratory stages of a regional

research program were described in Chapters I, II, and III. From these, a prioritized list of problems and promising new technology options will have been identified. The on-farm research team will have participated in those stages, and, in cooperation with commodity and discipline program leaders and station directors, can develop a work plan or research protocol in each research region. It may be useful to include specialists from other institutions in the analysis and planning sessions to broaden the overall viewpoint. The protocol should spell out in as much detail as possible the purpose and objectives of each regional on-farm research subprogram; the number and kinds of trials, experiments, or tests to be carried out; the methods and analyses to be used; the equipment, materials, and human and financial resources required; and the persons, commodity and disciplinary teams, or other services cooperating in the execution of the regional program and their contribution.

As a regional program continues over several years, the size and composition of the core team and the mix of needed back-stopping services will change. Initially, the geographical area for which the team is responsible should be kept small. The level of effort will be comparatively greater in the first years, as data needed to understand socioeconomic, biological, and technical problems are collected and analyzed. However, as the initial on-farm trials are completed and agro-climatic parameters and recommendation domains are refined and described, it may then be possible to increase the size of the area or reduce the size of the core team.

Each member of the team can be given responsibility for all research trials and farm records for a given set of farms even when these cut across several specialties. As noted, support and guidance can be provided by station-based specialists or more senior team members on a periodic basis. This method provides a built-in mechanism to improve the technical knowledge and research capability of each team member and, most importantly, the understanding of the farm as part of a complex biological and management system. The feedback of information to traditional research and academic institutions contributes to the orientation of basic research as well. Alternatively, depending on the size and complexity of the region, work may be divided within the team in part along

disciplinary lines. As an organizational framework, however, this alternative is usually less productive.

The specific number of trials and other activities one person can manage will depend, among other things, upon ease of access and mode of transportation available, distance between trial sites, the number of plots, and measurements and observations to be recorded. A typical workload for each team member might involve the supervision of 5 to 15 farmer-managed trials, responsibility for 8 to 12 site-specific or regional trials, and collection of crop records from 5 to 10 farms.

Off-station research also requires a different mix of research support staff, facilities, and equipment. Regional on-farm research programs succeed or fail on the strength and reliability of transportation and the flow of financial support for fuel, materials, salaries, and the like. Mobility of the team and timeliness in carrying out the work are critical for accurate, reliable research results and to maintain credibility with the farmers. A vehicle which will not run or seeds which arrive too late are of no immediate value to the research team and may cause serious delays in programs or financial loss to collaborating farmers.

DEFINING RESEARCH REGIONS

The ideal number of agricultural research regions defined for a country depends upon two main factors: 1) ecological diversity, and 2) the spatial distribution of farm system types. In practice, geographic limits of each research region are typically established in either of two ways. The most common is along geopolitical divisions of the country. Alternatively, agro-ecological boundaries may be used if they are known. The advantage of the first is that public services and planning and budget responsibilities are on the basis of political boundaries. The advantage of the agro-ecological approach is that major crop production zones, soil type boundaries, and farming system boundaries are often closely correlated. From a research management standpoint, this allows better coordination between on-station and on-farm research, better scientific support for regional teams, and easier interaction with the parent research organization. In either case the definition of recommendation domains

should guide the FSR/E activities within each region.

Factors which influence the choice of research regions where work is to begin or to be concentrated can include: the number of small, limited-resource farms; rural population density; the importance of the region for food and cash crop production; the agricultural potential of the region; and the potential to extrapolate research findings to similar regions. As a practical matter, the choice of research regions may be to support political objectives, such as counteracting rural unemployment or low income levels or enhancing a national development plan. In determining the number of research regions to which on-farm research teams are to be assigned, the number of separate research subjects, number of trials, and the like should be considered. The main criteria will be those that reach an equilibrium between priority needs and what is possible, given available resources. The importance of program planning and budgeting cannot be overstated. Once research is underway in several regions, interregional coordination must assure that information is exchanged and resources are allocated to each region and client group on an equitable and efficient basis.

EVALUATING ON-FARM RESEARCH

As research moves off station, some support and operation costs will be incurred. The logistics of the support and budgeting for on-farm research is an integral part of the planning process. These costs will be more than offset by the benefits of conducting more of the research under the physical, ecological, and socioeconomic conditions faced by farmers. Further, as leaders of commodity and disciplinary groups are able to carry out more trials on-farm, they are also better able to target on-station research toward the needs of client groups. Three principal groups of clients use the results of agricultural sector research. These are: farm families; agricultural sector support services which provide inputs and/or marketing services to farms; and public planning and policy clients, such as legislatures, planning departments, or financial ministries.

As an FSR/E program matures, some mechanism is needed by which both on-farm and on-station research activities can be evaluated, along with their contribution toward

reducing the effects of constraints that limit production. In this way, unproductive lines of investigation can be discarded earlier. A rigorous review procedure can help guard against a tendency of many scientists to pursue familiar research projects with only minor changes, even if these may no longer be yielding significant new information of potential use by clients.

An overall review will help identify where cooperative research or an exchange of information should take place. It will also identify duplication of effort or where resources could be better used on other priorities or in other regions. Another benefit of evaluation reviews is that participants in the process (at all levels, including support staff) develop a mission-oriented outlook and the overall program becomes more purposeful and rewarding.

A further advantage of FSR/E is that it helps to make the process of research and development more understandable to public leaders, farm client groups, and service organizations. It provides a visible example of public and institutional commitment to finding solutions for real problems and to improving rural life. These in turn can lead directly to public support for research and development, which may previously have been lacking, and indirectly to changes (where desirable) in public policies toward agriculture.

While it is rarely possible to measure the cost and benefit ratio of an FSR/E research program directly, it is possible to evaluate progress against goals and objectives, if these have been clearly stated. It may also be necessary to include measures or indicators of progress in data gathering and analysis activities. For example, the rate of adoption of technology introduced under the program is one indicator of progress, as are the net farm incomes, levels of production per unit area, or total production from an area measured over time. Readers familiar with the complexity of measuring such indicators, particularly under multiple cropping systems, will recognize the problem of objectively measuring program success.

Adding to the complexity, both research and farm managers must take into consideration the consequences of recommendations and decisions, not only for the current cropping season but also for future seasons. For example,

practices which control soil erosion, preserve soil moisture, and improve soil structure and fertility may reduce current earnings but also may improve future income and family security. Similarly, the planting of perennial crops, including trees for fuel and timber production, may reduce short-term profits but substantially increase future earnings. Conversely, practices or crops which mine soil fertility may increase current profits but endanger future earnings.

The rate at which new technology is adopted will depend foremost upon its acceptability to farmers. It also depends upon the effectiveness of government systems in providing the necessary support infrastructure, including roads, inputs to the production process, access to credit, and markets for products which are produced.

REGIONAL INTERINSTITUTIONAL COOPERATION

The International Agricultural Research Centers and larger regional centers, such as CATIE in Central America, have played and will continue to play a key role in support of national research programs, particularly in three areas: 1) germ plasm collection and advanced breeding; 2) training of scientists, commodity specialists, and technical staff; and 3) exchange of technical information through publications, special subject seminars, and exchange visits among scientists and other specialists. Many of the international and regional centers have been active in developing FSR methodology and networks because of complementarity with their program objectives. This activity constitutes a natural link with national programs, which are best able to take advantage of the resources provided by the centers in the way of material, training, and information.

Other public and private enterprises which either have a direct interest in the information flowing out of an FSR/E program or could provide information and materials helpful to the FSR/E process include:

a) agencies responsible for national economic surveys, development plans, farm surveys, or censuses;
b) enterprises producing or marketing improved seeds, pesticides, and farm equipment;
c) farm credit agencies which provide short-term

production credit and loans for capital improvements;

d) commodity marketing services and companies, including cooperatives interested in expanding sources of supplies;

e) public and private laboratories which provide soil testing services and disease diagnosis; and

f) formal and non-formal education and training institutions desiring to make training and materials more relevant to students.

Well-designed and implemented on-farm research programs are attractive to policy makers and infrastructure managers because it is easier for them to identify with the research in progress. As information becomes more specific by groups of farms, it becomes easier to target service development project and training to each group of clients. In this way, it becomes easier to conform research and development to national or regional policy goals and objectives, such as small farm development or food sufficiency goals.

Bibliography

Byerlee, D., M. Collinson, et al. 1980. Planning technologies appropriate to farmers--concepts and procedures. CIMMYT, Mexico.

Chew, Victor. 1977. Comparisons among treatment means in an analysis of variance. Agricultural Research Service, U.S. Department of Agriculture, Technical Bulletin # ARS/H/6.

Cochran, W. G., and G. M. Cox. 1957. Experimental designs. 2nd edition. John Wiley and Sons, Inc., New York.

Duarte, R., P. E. Hildebrand, and S. R. Ruano. 1977. Tecnologia y estructura agro-socioeconomica del minifundio del occidente de Chimaltenango. Socioeconomia Rural. ICTA, Guatemala.

Fox, Karl. 1968. Intermediate economic statistics. John Wiley and Sons, Inc., New York.

Previous Page Blank

Hansen, A., E. N. Mwango, and B. S. C. Phiri. 1982. Farming systems research in Phalombe project, Malawi: another approach to small holder research and development. Center for Tropical Agriculture, University of Florida. In cooperation with Farming Systems Analysis Section, Department of Agricultural Research, Ministry of Agriculture, Government of Malawi.

Hart, R. D. 1979. Agroecosistema: Conceptos basicos. Centro Agronomico Tropical de Investigacion y Enseñanza, Turrialba, Costa Rica.

Hildebrand, P. E. 1972. Analisis agroeconomicos mediante superficies de respuesta. Ministerio de Agricultura y Ganaderia, Direccion General de Economia Agricola y Planificacion, Departamento de Administracion Agricola, El Salvador.

Hildebrand, P. E. 1984. Modified stability analysis of farmer managed, on-farm trials. Agron J. 76:271-274.

Hildebrand, P. E., S. R. Ruano, T. Lopez Yos, E. Samayoa, and R. Duarte. 1977. Sistemas de cultivos para los agricultores tradicionales del occidente de Chimaltenango. Socioeconomia Rural. ICTA, Guatemala.

Hildebrand, P. E., and R. K. Waugh. 1983. Farming systems research and development. Farming Systems Support Project Newsletter, Vol. 1, No. 1. University of Florida, Gainesville.

Little, T., and F. Jackson Hills. 1978. Agricultural experimentation: design and analysis. John Wiley and Sons, Inc., New York.

Montgomery, Douglas C. 1976. Design and analysis of experiments. John Wiley and Sons, Inc., New York.

Norman, D. W. 1982. The farming systems approach to research. Farming Systems Research Paper No. 3. Kansas State University, Manhattan, Kansas.

Ruano, Sergio R. 1981. Evaluacion de la aceptabilidad de la tecnologia generada por el ICTA en arroz en el area de Chiquimulilla. Socioeconomia Rural. ICTA, Guatemala.

Samayoa, E. J. 1980. Evaluacion de aceptabilidad de la tecnologia generada por el ICTA para los sistemas de surco doble de maiz con calles anchas, Chimaltenango, 1979-1980. Socioeconomia Rural. ICTA, Guatemala.

Shaner, W. W., P. F. Philipp, and W. R. Schmehl. 1982. Farming systems research and development: guidelines for developing countries. Westview Press, Boulder, Colorado.

Snedecor, G. W. 1956. Statistical methods. The Iowa State College Press, Ames, Iowa.

Steel, R. G. D., and J. H. Torrie. 1980. Principles and procedures of statistics. 2nd edition. McGraw-Hill Book Company, New York.

Zandstra, H. G., E. C. Price, J. A. Litsinger, and R. A. Morris. 1981. A methodology for on-farm cropping systems research. The International Rice Research Institute. Los Baños, Philippines.

About the Book and Authors

The authors have synthesized and elaborated on the work of a team of agricultural and social scientists expert in farming systems research and extension (FSR/E). Their handbook presents the role and philosophy of on-farm research in FSR/E activities, describes a logical sequence for technology development, and explains the most-used statistical procedures in easy-to-follow steps. It also outlines new ideas and methods for analyzing agronomic data obtained outside of the controlled conditions typically found on experiment stations.

Peter E. Hildebrand is professor in the Food and Resource Economics Department at the University of Florida. Previously, he was with the Rockefeller Foundation and on the faculties of the University of Nebraska, Colorado State University, and Texas A & M University. He has been involved in development projects in Pakistan, Colombia, El Salvador, and Guatemala. **Federico Poey**, president of Agricultural Development Consultants, Inc., has worked as a seed specialist with CIAT and a maize breeder and agronomist with CIMMYT. From 1954 to 1976 Dr. Poey was breeding and production director for Semillas Poey in Cuba and Mexico and then maize breeding director (for tropical America) for Northrup King Company.