Statistical Design and Analysis for Intercropping Experiments, Volume II: Three or More Crops

Walter T. Federer

Springer

Springer Series in Statistics

Springer
New York
Berlin
Heidelberg
Barcelona
Budapest
Hong Kong
London
Milan
Paris
Singapore
Tokyo

Springer Series in Statistics

Andersen/Borgan/Gill/Keiding: Statistical Models Based on Counting Processes.
Andrews/Herzberg: Data: A Collection of Problems from Many Fields for the Student and Research Worker.
Anscombe: Computing in Statistical Science through APL.
Berger: Statistical Decision Theory and Bayesian Analysis, 2nd edition.
Bolfarine/Zacks: Prediction Theory for Finite Populations.
Borg/Groenen: Modern Multidimensional Scaling: Theory and Applications
Brémaud: Point Processes and Queues: Martingale Dynamics.
Brockwell/Davis: Time Series: Theory and Methods, 2nd edition.
Daley/Vere-Jones: An Introduction to the Theory of Point Processes.
Dzhaparidze: Parameter Estimation and Hypothesis Testing in Spectral Analysis of Stationary Time Series.
Fahrmeir/Tutz: Multivariate Statistical Modelling Based on Generalized Linear Models.
Farrell: Multivariate Calculation.
Federer: Statistical Design and Analysis for Intercropping Experiments, Volume I: Two Crops.
Federer: Statistical Design and Analysis for Intercropping Experiments, Volume II: Three or More Crops.
Fienberg/Hoaglin/Kruskal/Tanur (Eds.): A Statistical Model: Frederick Mosteller's Contributions to Statistics, Science and Public Policy.
Fisher/Sen: The Collected Works of Wassily Hoeffding.
Good: Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses.
Goodman/Kruskal: Measures of Association for Cross Classifications.
Gouriéroux: ARCH Models and Financial Applications.
Grandell: Aspects of Risk Theory.
Haberman: Advanced Statistics, Volume I: Description of Populations.
Hall: The Bootstrap and Edgeworth Expansion.
Härdle: Smoothing Techniques: With Implementation in S.
Hart: Nonparametric Smoothing and Lack-of-Fit Tests.
Hartigan: Bayes Theory.
Heyde: Quasi-Likelihood And Its Application: A General Approach to Optimal Parameter Estimation.
Heyer: Theory of Statistical Experiments.
Huet/Bouvier/Gruet/Jolivet: Statistical Tools for Nonlinear Regression: A Practical Guide with S-PLUS Examples.
Jolliffe: Principal Component Analysis.
Kolen/Brennan: Test Equating: Methods and Practices.
Kotz/Johnson (Eds.): Breakthroughs in Statistics Volume I.
Kotz/Johnson (Eds.): Breakthroughs in Statistics Volume II.
Kotz/Johnson (Eds.): Breakthroughs in Statistics Volume III.
Kres: Statistical Tables for Multivariate Analysis.
Küchler/Sørensen: Exponential Families of Stochastic Processes.
Le Cam: Asymptotic Methods in Statistical Decision Theory.

Springer Series in Statistics

(continued from p. ii)

Le Cam/Yang: Asymptotics in Statistics: Some Basic Concepts.
Longford: Models for Uncertainty in Educational Testing.
Manoukian: Modern Concepts and Theorems of Mathematical Statistics.
Miller, Jr.: Simultaneous Statistical Inference, 2nd edition.
Mosteller/Wallace: Applied Bayesian and Classical Inference: The Case of The Federalist Papers.
Parzen/Tanabe/Kitagawa: Selected Papers of Hirotugu Akaike.
Pollard: Convergence of Stochastic Processes.
Pratt/Gibbons: Concepts of Nonparametric Theory.
Ramsay/Silverman: Functional Data Analysis.
Read/Cressie: Goodness-of-Fit Statistics for Discrete Multivariate Data.
Reinsel: Elements of Multivariate Time Series Analysis, 2nd edition.
Reiss: A Course on Point Processes.
Reiss: Approximate Distributions of Order Statistics: With Applications to Non-parametric Statistics.
Rieder: Robust Asymptotic Statistics.
Rosenbaum: Observational Studies.
Ross: Nonlinear Estimation.
Sachs: Applied Statistics: A Handbook of Techniques, 2nd edition.
Särndal/Swensson/Wretman: Model Assisted Survey Sampling.
Schervish: Theory of Statistics.
Seneta: Non-Negative Matrices and Markov Chains, 2nd edition.
Shao/Tu: The Jackknife and Bootstrap.
Siegmund: Sequential Analysis: Tests and Confidence Intervals.
Simonoff: Smoothing Methods in Statistics.
Small: The Statistical Theory of Shape.
Tanner: Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions, 3rd edition.
Tong: The Multivariate Normal Distribution.
van der Vaart/Wellner: Weak Convergence and Empirical Processes: With Applications to Statistics.
Vapnik: Estimation of Dependences Based on Empirical Data.
Weerahandi: Exact Statistical Methods for Data Analysis.
West/Harrison: Bayesian Forecasting and Dynamic Models, 2nd edition.
Wolter: Introduction to Variance Estimation.
Yaglom: Correlation Theory of Stationary and Related Random Functions I: Basic Results.
Yaglom: Correlation Theory of Stationary and Related Random Functions II: Supplementary Notes and References.

Walter T. Federer

Statistical Design and Analysis for Intercropping Experiments

Volume II: Three or More Crops

With 34 Illustrations

Springer

Walter T. Federer
Biometrics Unit
Cornell University
Ithaca, NY 14853-7801
USA

Library of Congress Cataloging-in-Publication Data
Federer, Walter Theodore, 1915–
Statistical design and analysis for intercropping experiments by Walter T. Federer.
p. cm. — (Springer series in statistics)
Includes bibliographical references and index.
Contents: v. 1. Two crops
ISBN 0-387-98533-6
1. Intercropping—Experiments. 2. Experimental design.
I. Title. II. Series.
S603.5.F43 1993
631.5′8—dc20 92-29586

ISBN 0-387-98533-6 Springer-Verlag New York Berlin Heidelberg SPIN 10659500

To Edna

Preface to Volume II

The comments about Volume II in the Preface to Volume I were overoptimistic. In order to obtain the statistical designs and analyses, it was necessary to perform an inordinate amount of research to prepare new and more complex treatment and experiment designs and associated statistical analyses. A considerable amount of algebraic manipulation was required. Owing to the algebraic complexity, it was useful to employ computer software such as GAUSS, MATHEMATICA, and MAPLE. Specific cases were obtained using these packages, which enabled a generalization. Some of the programs are included to aid the reader in developing computer programs for other situations. As intercropping research is a complex and broad field, many new situations, which were not contemplated when Volume II was written, will arise. Use of the above packages resulted in saving a considerable amount of time in arriving at solutions. Estimation of other effects and their variances will be needed and other goals will arise in meeting the needs of future intercropping research.

The scope of Volume II is given in the first chapter. The simplest case of intercropping with three or more crops is one main crop with k supplementary crops as discussed in Chapter 12. The statistical analyses are much more complicated than when there is only one supplementary crop. The complexity of the statistical design and analysis continues for all other situations. From a study of the literature citations in the bibliography described in Chapter 20, intercropping research has been concerned almost entirely with rather simple intercropping goals, designs, and analysis. With the designs and analyses presented herein, it will be possible for the researchers to study more complex, meaningful, and practical situations. The past studies, for a large part, deal with only one component of the agricultural cropping system. Analyses given herein allow for a consideration of the entire system as well as its component parts.

A comparison of systems or programs arises in many fields besides agriculture. Even in agriculture there are many systems involving mixtures of items which are not ordinarily thought of as intercropping, but in reality are. Some examples are pasture, orchard, lumber plantations, tea and coffee plantations, weed control, rotation, and double cropping studies as described in Chapter 19. In medicine, patients use many types of drug mixtures, sometimes taking ten or more prescription drugs. Drug "cocktails" are being prescribed. Many of the designs and analyses presented here may be used in studying the various effects of drug mixtures. In education and recreation, many programs are used and the procedures given will be useful in assessing the various programs. Mixtures of items appear in engineering, transportation, nutrition, ecology, etc. studies. Procedures for intercropping studies have direct application in all these areas.

Walter T. Federer, 1998

Preface to Volume I

Two volumes are being published on the topic of the title. Volume I, the present one, deals with the statistical design and analysis of intercropping experiments in which there are mixtures (intercrops) of two crops and/or sole crops. Volume II will deal with the statistical design and analysis of three or more crops in the mixture (intercrop), together with sole crops and possible mixtures of two crops. It is necessary to fully comprehend the concepts and analyses for mixtures of two crops prior to considering three or more crops in the mixture. The utility, concepts, comprehensions, and application of techniques for two crops are an order of magnitude more difficult than for sole crops only. The degree of difficulty in these aspects for three or more crops in the mixture is an order of magnitude greater than when considering only two crops. Hence, the reader is cautioned to fully comprehend Volume I before proceeding to Volume II. Most published literature deals with two crops in a mixture. In practice, the number of crops in an intercropping system may be quite large. Mixtures of three or more crops are quite common in practice, e.g., pastures. The last chapter of Volume I considers design concepts and experiment designs that may be of use for intercropping experiments. The last two chapters of Volume II will contain a bibliography of publications on intercropping, which are not cited at the end of each chapter in Volumes I and II, and a discussion of applications of the material for intercropping experiments to other areas. Some of the areas are survey sampling, chemistry, hay crop mixtures, repeated block designs, dietary studies, and recreational and educational programs.

In presenting the statistical design and analysis for intercropping experiments, involving mixtures of two crops with or without sole crops, we have attempted to present the topics in order of increasing difficulty. First, the situation involving one main crop and one supplementary crop is considered in Chapter 2. Here, we add little over that appearing in standard statistical methods books. Then, in Chapter 3, we consider both crops to be main crops and analyze the individual crop responses. Again, little is added on statistical methodology that is not standard. In Chapter 4, both crops are considered to be main crops and a combined response for the yields of both crops is required. This involves creating variables as in a multivariate analysis. The forms used will not ordinarily be those from standard multivariate analyses. Ratios of yields, prices, or other variables are used. This is an innovation over other procedures appearing in the literature. We show that several analyses are desirable, as opposed to one when only sole crops are in the experiment. Density of crops is held constant up to here. In Chapter 5, density is a variable for the two main crops and yield is modeled as a function of density. In Chapter 6, we model responses in much the same way as they are for diallel crossing systems in breeding investigations, *except* that the yields from both crops are available. In Chapter 7, we do the same type of modeling for the case when the individual crop yields are *not* available. This is closer to the ordinary diallel crossing situation. In Chapter 8, spatial arrangements of two crops are discussed, with many arrangements being considered. In the ninth chapter of Volume I, analyses of replacement series experiments and a linear programming approach for considering two responses simultaneously in a replacement series are discussed. The last chapter contains a discussion of design concepts and experiment designs that are considered to be of use in intercropping experiments.

As of this date, most of the theoretical work for Volume II, Chapters 11 to 20, has been completed. Chapters 12 and 13 have already been written and are in the process of being put in final form. A search for appropriate examples is being made. A bibliography on intercropping experiments, Chapter 20, has been made but will require updating.

WALTER T. FEDERER, 1993

Acknowledgments

The name of Nancy Miles McDermott was inadvertently omitted in the acknowledgments to Volume I. Deep appreciation and many thanks are due to her for her efforts in assuring correctness of the computations and for her enlightening comments. It was a joy to work with her.

Norma Phalen's effort in preparing the manuscript for this book in LaTex format is greatly appreciated. Thanks are due her for reading over the manuscript, checking for spelling errors, making comments, and especially her patience with the entire process. The efforts and suggestions of the Senior Production Editor, Lesley Poliner, and Copyeditor, Barbara Tomkins, are sincerely appreciated. Ladies, I couldn't have done it without you. Many, many thanks.

WALTER T. FEDERER, 1998

Contents

List of Tables

List of Figures

CHAPTER 11

Introduction to Volume II

11.1 Experiments Involving Comparisons of Agricultural Systems

Experiments involving the comparison of systems, such as agricultural systems, medical treatments systems, educational systems, etc., require a multi-faceted approach for setting up the goals of the investigation, in designing the experiment, and in performing the necessary statistical analyses. (See, e.g., Kass, 1978, Mead and Riley, 1981, Balaam, 1986, Federer, 1987, 1989, 1993a, 1993b, hereafter referred to as Volume I, and references therein.) When performing experiments comparing agricultural systems, the researcher needs to consider goals involving efficiency of land use, nutritional values, economic values, sustainablity of yields in the system, insect and disease control, soil structure and erosion, spatial arrangements of the system, density and intimacy considerations, competition, mixing abilities of components of the system, and/or perhaps other characteristics. In most cases, it is not be possible to generalize from monocultures to polycultures, from pairs of cultivars to mixtures of more than two, and so forth. Four rules to keep in mind when conducting intercropping experiments are as follows:

Rule 1. Understand the concepts, design, and analyses for mixtures of two crops before proceeding to mixtures of three or more cultivars.

Rule 2. Do not attempt to generalize from monocultures to pairs of cultivars, from pairs to triplets of cultivars, from triplets to quartets of cultivars, from one set of cultivars to another, and so on, as this may lead to gross errors.

Experiments need to be conducted for the specific mixture size and the specific cultivars under study.

Rule 3. Be prepared for the increasing difficulty of design, analysis, and interpretation involved, as the degree of difficulty increases by an order of magnitude in going from monocrops to mixtures of two, by another order of magnitude in going from pairs to triplets, etc.

Rule 4. Be prepared for and look for surprises, as many intercropping experiments produce quite unexpected results, as was exhibited in the examples in Volume I and in the examples presented herein.

Intercropping is an age-old practice going back at least to early Biblical times (The Holy Bible, 1952). It is a farming system that is popular in many areas of planet Earth, especially in tropical agriculture but is present in some form all over the world. Even in temperate zone agriculture, intercropping is common in hay crops, in orchard cover crops, in crop rotations, and in cover crops for such crops as alfalfa. Many gardeners use crop mixtures and sequences for a variety of reasons, one being insect and disease control. In making comparisons among agricultural systems, a variety of statistical designs and analyses will be required and will be demonstrated in the following chapters. But first let us consider some of the goals, uses, and other considertions of intercropping systems investigations.

11.2 Land Use and Agronomic Goals

As Earth's populations tend to increase and with agricultural land area being depleted by urbanization and salinization, it is necessary to make more and more efficient use of the available agricultural land area. A measure for efficiency of land use is the relative yield (de Wit and van den Bergh, 1965) or land equivalent ratio (Willey and Osiru, 1972). A land equivalent ratio (LER) is an agronomic characteristic of an intercropping experiment. It is the sum of ratios of yields of a crop, say i, in a mixture, say Y_{mi}, to its yield as a sole crop, say Y_{si}. Then, for n crops, an LER is

$$\mathrm{LER} = \sum_{i=1}^{n} Y_{mi}/Y_{si} = \sum_{i=1}^{n} \mathrm{LER}_i . \tag{11.1}$$

Instead of using the yield of the sole as the denominator, another form of an LER could be obtained by using the yield of crop i in a standard mixture. A variety of other values could be used for Y_{si} in (11.1) such as

- individual plot yields of the sole crop,
- mean yields from r replicates for the sole crop,
- a theoretical "optimum value" for the sole crop,
- farmer's yields averaged over y years or for a single year for the sole crop, or
- some other value.

As is obvious, there are many possible LERs. Therefore, it is imperative that the experimenter understands the properties of and consequences of using the LER selected for the determination of land use efficiency of a cropping system.

When only the numerators in an LER are random variables and the denominators are fixed constants, then standard statistical procedures are available for use as explained in Volume I. When both numerators and denominators in the LER are random variables, little is known about the statistical distribution of the LERs. If the numerators and denominators are random normal deviates from a multivariate normal distribution, then the statistic in (11.1) has a Cauchy distribution (Federer and Schwager, 1982) which has infinite variance. If the numerators and denominators come from log-normal distributions, Morales (1993) has obtained the statistical distribution for two crops in the mixture. Presently, work is being done considering the distributions of sums of ratios of gamma-distributed random variables, but at this writing, this research is not at the stage of practical usefulness. A normal distribution ranges from plus to minus infinity. Hence, crop responses not having this range as a possibility cannot be normally distributed. Gamma random variables range from zero to plus infinity, which has a realistic starting point, zero, for yield, counts, etc.

As described in Volume I, one way out of this dilemma is when one sole crop can be used as a base sole crop, say Y_{s1}. Then, use ratios of yields of sole crops to the base sole crop, say crop 1, as follows to obtain a relative land equivalent ratio (RLER):

$$\text{RLER} = \sum_{i=1}^{n} Y_{s1} Y_{mi} / Y_{si} = \sum_{i=1}^{n} R_i Y_{mi} = \sum_{i=1}^{n} \text{RLER}_i . \tag{11.2}$$

A RLER is useful in comparing cropping systems and statistical analyses but needs to be converted to an LER for actual land-use considerations. Ratios of yields and prices, e.g., are much more stable than are actual yields and prices (Ezumah and Federer, 1991). Since this is true, the ratios R_i may be regarded as fixed constants rather than as random variables, and the problem of the distribution of sums of ratios of random variables is bypassed to one which is simply a linear combination of random variables.

11.3 Crop Value and Economic Goals

Various values may be assigned to the yield of each crop in a mixture. For many people, value means monetary value. For others, value could be related to how well dietary goals of a family are satisfied with regard to taste and variety of foods in a diet. Crop value for others could be related to frequency of produce for sale or barter throughout the year. Whatever value system is used, consider the value, monetary or otherwise, of crop i to be P_i per crop unit, such as a kilogram or individual fruit. The value of a crop will then be $P_i Y_{mi}$, where Y_{mi} is the total yield or number of fruit per experimental unit (e.u.). Then, the value of the crops in a mixture of n crops is

$$\text{Crop value} = V = \sum_{i=1}^{n} P_i Y_{mi}. \tag{11.3}$$

Although prices or other crop values may fluctuate considerably from year to year, ratios of prices or values may not (Ezumah and Federer, 1991). Hence, for comparative purposes, relative crop values may be used and the difficulties of random fluctuations in prices avoided. As for RLER, a base crop price is selected, say P_1, and ratios of crop values are used to obtain a relative crop value, RV, for a mixture of n crops as

$$\text{RV} = \sum_{i=1}^{n} (P_i/P_1) Y_{mi}. \tag{11.4}$$

The goal would be to select that mixture maximizing V or, equivalently, RV. In making comparisons of the v mixtures in an experiment, it is recommended that RLER and RV be utilized in order to circumvent statistical distribution problems. Their use will also ease presentation problems of the several analyses required to summarize the information from intercropping experiments.

11.4 Nutritional Goals

In subsistence farming areas of the world, the number of calories provided by the crops grown on the farm is of vital importance. Insufficient calories in the diet leads to dietary difficulties and to starvation in extreme cases. Protein content is also important for a proper diet. Palatability of the foods produced is of concern, as it will not matter how many calories are produced if the produce is unpalatable and cannot be used for sale or barter. In intercropping experiments, it is necessary to assess the caloric and protein content of mixtures and sole crops and the palatability of the foods produced.

For comparative purposes, calorie conversion factors for the various crops in a mixture are available. These conversion factors may vary widely between crops and less so among cultivars within crops. After selection of appropriate conversion factors for each of the crop cultivars in the mixture of n cultivars, the total calories, protein, or other measure is

$$C = \sum_{i=1}^{n} C_i Y_{mi}, \tag{11.5}$$

where C_i is the conversion factor for cultivar i and for the characteristic under consideration. A relative total calorie, total protein, total fiber, total vitamin, etc., for crop 1 as the base crop is

$$\text{RC} = \sum_{i=1}^{n} C_i Y_{mi}/C_1 = \sum_{i=1}^{n} R_i Y_{mi}. \tag{11.6}$$

An RV or RC may not appear appropriate at first glance, but using only relative measures RLER, RV, and RC affords ease of presentation of the several analyses used, e.g., putting results on the same graph.

Note that the C_i in (11.5) and (11.6) could be of a complex form if it were decided to combine nutritional measures. Suppose the relative importance of protein conversion factor C_{pi} to the carbohydrate conversion factor C_{ci} is $R_{p/ci}$, of the fiber conversion C_{fi} to carbohydrate is $R_{f/ci}$, of the vitamin conversion factor C_{vi} to carbohydrate is $R_{v/ci}$, etc., then the conversion factor for all components could be of the form

$$C_i Y_{mi}/C_{ci} = (1 + R_{p/ci} + R_{f/ci} + R_{v/ci} + \cdots) Y_{mi}, \tag{11.7}$$

where C_i in (11.5) and (11.7) is equal to

$$C_i = C_{ci} + C_{pi} + C_{fi} + C_{vi} + \cdots. \tag{11.8}$$

This form of C_i could be used in (11.5) and (11.6). Also, different weights could be added to take into account the relative importance of carbohydrates, protein, fiber, vitamins, and other dietary components as a measure of the nutritional value of a mixture.

11.5 Sustainability of a System

The term sustainability has many and diverse meanings in published literature. Therefore, it behooves the author to state which definition is being used. For example, does sustainability mean

- constant crop yields year after year,
- fluctuations in yearly yields but no downward or upward trends in yield,
- the above two situations but crop value replacing yield,
- a system that has survived through time, or
- yield to meet population nutritional requirements over time?

Or does it follow the definition

> A sustainable agriculture is one that, over the long term, enhances environmental quality and the resource base on which agriculture depends; provides for basic human food and fiber needs; is economically viable; and enhances the quality of life for farmers and society as a whole. (Anon., 1989)

Does it follow the definition in the 1990 Farm Bill which mandated the USDA to support research and extension in sustainable agriculture defined as

> An integrated system of plant and animal production practices having a site-specific application that will over the long term: (i) satisfy human and fiber needs; (ii) enhance environmental quality and the natural resources base upon which the agricultural economy depends; (iii) make the most efficient use of nonrenewable resources and on-farm resources, and integrate,

> where appropriate, the natural biological cycles and controls; (iv) sustain the economic viability of farm operations; and (v) enhance the quality of life for farmers and society as a whole.

Weil (1990) prefers the following definition

> An agricultural program, policy, or practice contributes to agricultural sustainability if it:
>
> **1.** Enhances, or maintains, the number, quality, and long-term economic viability of farming and other agricultural business opportunities in a community or region.
> **2.** Enhances, rather than diminishes, the integrity, diversity, and long-term productivity of both the managed agricultural ecosystem and the surrounding ecosystems.
> **3.** Enhances, rather than threatens, the health, safety, and aesthetic satisfaction of agricultural producers and consumers alike.

In each of the last three definitions, there are several words or phrases which could be interpreted in several ways. There are several undefined terms. This leads to the conclusion that a clear, precise, unambiguous, and meaningful definition of the term sustainablity still needs to be discovered.

Despite the multiplicity of interpretations possible for the term sustainability, an agricultural system is a sustainable one if it has endured the ravages of time. Intercropping, in its broadest sense as defined in Volume I, is one practice that has endured since early times. Rotational cropping, one form of intercropping, has been used for centuries to control erosion, enhance crop yields, control disease and insects, and provide a variety of crops. Whether sequentially or simultaneously growing mixtures of several crops, this farming system has endured and, hence, was sustainable. Aina *et al*. (1977), Lal (1989), Ezumah and Hullugalle (1989), and Hullugalle and Ezumah (1989) have demonstrated experimentally that intercropping results in better soil structure and less erosion than sole cropping. As pointed out by Federer (1989), the use of chemicals in the developed countries has created a "chemical agriculture" to replace the tried and true long-term agricultural system of intercropping. It has yet to be determined if chemical agriculture will pollute our water supplies, destroy wildlife, and cause human sickness over the long term. Evidence is mounting that chemical fertilizers and pesticides should be used sparingly, if at all. The disastrous effect of using DDT is well known and documented. Agricultural systems are available for which no chemical fertilizer is required, e.g., an intercropping system involving *Leucaena leucocephalia*, maize, and beans, a rotational system involving legumes and grass species, and composting and gardening. Chemical agriculture is not necessary; it is simply convenient and economically viable for the present. It may not be economically viable when the cost of cleaning up the environment and/or paying medical bills is added to the cost of production. Also, it is likely that the cost of chemicals will increase as energy supplies decrease or become more costly and make chemical agriculture economically nonviable (Pimentel *et al*., 1994). A return to centuries-old agricultural systems, or some modification of them, may be in the future for agricultural production.

11.6 Biological Goals and Considerations

In addition to agronomic, economic, and nutritional considerations in analyzing data from an intercropping experiments, it is often important to determine the nature of biological phenomena involved in intercropping systems. The determination and measurement of how well cultivars mix or compete, how mixtures respond to density changes and spatial arrangements and why, synergistic relationships and mechanisms, and possibly new biological concepts are some of the biological considerations required when interpreting the results from an intercropping experiment. Yield–density relationships need to be modeled. Measures of mixing ability need to be developed. Competition models for various situations need to be available. Knowledge of the biological processes governing the responses of why some systems or mixtures perform as they do is necessary in order to develop methods for producing the desirable systems or mixtures in an efficient manner. Knowing the theory behind a system is helpful to the researcher in producing a more desirable system. This situation has precedence in plant breeding where diallel crossing, top-crossing, single-crossing, double-crossing, and multiple-crossing theory and procedures were developed and applied to develop the desired cultivars. The concepts and results of Chapters 5, 6, and 7 in Volume I are extended to mixtures of more than two cultivars in the following chapters.

11.7 Statistical Considerations

The topic of experiment design, the arrangement of treatments in an experiment, has been covered in Chapter 10 of Volume I. The experiment design is for v treatments for whatever treatment design is used. The control of experimental heterogeneity by blocking or covariance is a topic independent of the treatments included in an experiment. Treatment design, the selection of treatments to be used in an experiment, in intercropping studies is vital in reaching desired goals. Since there are many goals and situations, there will be a variety of treatment designs. Since the number of treatments v in an intercropping experimment can become large quickly, it is necessary to select minimal treatment designs (TDs). Minimal TDs which contain as many treatments as there are independent parameters to estimate are called saturated designs. If all independent parameters are estimable, the TD is said to be connected. Thus, saturated designs which are minimal and connected are desired. TDs are needed for the situation where a response for each member of a mixture of n crops is available and when only one response is available for the mixture. As will be demonstrated in the following chapters, many and diverse TDs are required in intercropping investigations. Experiment design theory involving balanced incomplete block, partially balanced incomplete block, Youden, and supplemented block designs is utilized to construct the various and diverse TDs. Methods other than trial and error are needed for construction of some of the saturated TDs.

In general, any mixture of n crops of interest qualifies for inclusion in an intercropping investigation. For certain goals and analyses, it may be necessary to include sole crops and all possible mixtures of size n of m cultivars. The particular treatment design selected needs to be done considering the precisely defined goals of the experiment. If, e.g., the goal is to compare v mixtures with a standard sole crop or mixture, this is only possible when the standard or appropriate sole crop is included. Appropriate standards as points of reference should be included in the TD. In selecting a TD, the experimenter should consider the following rules:

1. Precisely define the goals of the investigation.
2. Select treatments allowing accomplishment of stated goals.
3. Consider the TD in light of the anticipated statistical analyses.
4. Decide in light of steps 1, 2, and 3 if the required comparisons are possible.
5. Revise steps 1, 2, and 3, if step 4 is not answered in the affirmative.

In place of conducting several small experiments, it may be possible to combine them into a single experiment creating many times the information obtained from the single experiments. By combining experiments, additional information may be available on the interaction of the treatments in the two experiments as well as comparing the treatments over a wider range of conditions. This situation arises frequently when consulting with researchers about their investigations. Combining several small experiments often leads to conservation of space, material, and labor, resulting in more efficient use of resources. Investigators from different fields often can use the same experimental material for their studies. For example, an entomologist, pathologist, and plant breeder can often use the same varietal experiment for their studies rather than conducting three separate varietal trials. Factorial treatment designs are more informative and efficient than experiments on the separate factors and can be used to combine the treatments from separate experiments.

Modeling yield–density relations for sole crops is much simpler than modeling yield–density relations for the n cultivars in the mixture. It is necessary to determine which, if any, of the cultivars in a mixture are to have density varied. Varying densities for all n crops in mixture will necessarily require many experimental units (e.u.s). Hence, the experimenter should only include enough densities to model the yield–density relationships. In addition to density considerations, spatial and intimacy (the nearness of cultivars in a mixture) of the n cultivars in the mixture need to be taken into account. Are the cultivars to be in separate rows, mixed together in the same row, some combination of the previous two, or to be in a broadcast arrangement? Are cultivars included in a mixture at different times? Is every cultivar bordered by every other cultivar and on one side, or all sides? These are items of importance in intercropping studies and require the attention of the intercrop researcher. Plot technic regarding shape and size of an e.u. is also important. From an experiment design standpoint, long narrow e.u.s over all types of gradients are more efficient than square e.u.s. For intercropping studies, long narrow e.u.s may be ineffective because of the intimacy, competition, and mixing ability characteristics required to evaluate a mixture to be used in practice.

The linear combination of responses for the n cultivars in a mixture discussed in Sections 11.2, 11.3, and 11.4 are reminicent of canonical variates in multivariate analyses. The statistician unfamilar with intercropping might think that multivariate statistical techniques would satisfy the needs of statistical analysis. However, as Federer and Murty (1987) have pointed out, multivariate techniques have very limited usefulness in this area. One use for mixtures of size two has been demonstrated by Pearce and Gilliver (1978, 1979). The multivariate analysis mathematical criterion used to a canonical variate is to select a linear combination of the responses for the n items, say,

$$\text{first canonical variate} = \sum_{i=1}^{n} a_i Y_{mi}, \tag{11.9}$$

in such a way that no other selection of the a_i has a larger ratio of the treatment sum of squares to the treatment sum of squares plus error sum of squares. Then, to the residuals from the first canonical variate, a second canonical variate, say, is constucted as

$$\text{second canonical variate} = \sum_{i=1}^{n} b_i Y_{mi}, \tag{11.10}$$

where the b_i are selected in the same manner as the a_i, and so forth, until n canonical variates are obtained. As pointed out by Federer and Murty (1987), the a_i and b_i have no practical interpretation and, hence, are of no use to the experimenter. These authors also describe other difficulties in trying to apply standard mutivariate techniques to the results from intercropping experiments.

11.8 Scope of Volume II

In Volume I, the following chapters were included:

Chapter 1. Introduction and Definitions
Chapter 2. One Main Crop Grown with a Supplementary Crop
Chapter 3. Two Main Crops—Density Constant—Analyses for Each Crop Separately
Chapter 4. Both Main Crops—Density Constant—Combined Crop Responses
Chapter 5. Both Crops of Major Interest with Varying Densities
Chapter 6. Monocultures and Their Pairwise Combinations When Responses Are Available for Each Member of the Combinations
Chapter 7. Monocultures and Their Pairwise Combinations When Separate Crop Responses Are Not Available

Chapter 8. Spatial and Density Arrangements
Chapter 9. Some Analytical Variations for Intercropping Studies
Chapter 10. Experiment Designs for Intercropping Experiments

In Volume II, the chapters are numbered consecutively following those in Volume I and these chapters, with their relationships to the chapters in Volume I, are as follows:

Chapter 11. Introduction to Volume II

• This chapter is a continuation of the material in Chapter 1 in Volume I to cover considerations, goals, and experimental objectives involved when more than two crops make up a mixture.

Chapter 12. One Main Crop Grown with More Than One Supplementary Crop

• The results in Chapter 2 are extended in this chapter to cover mixtures involving three or more crops in the mixture. As noted, extension of the ideas and procedures in Chapter 2 is not straightforward but require more sophisticated procedures.

Chapter 13. Three or More Main Crops—Density Constant

• The results of Chapters 3 and 4 of Volume I are extended in this chapter to cover mixtures involving three or more crops in the mixture. As noted, extension of the ideas and procedures in Chapters 3 and 4 are not straightforward but require more complicated procedures such as land equivalent ratios, relative land equivalent ratios, etc. for three or more crops in a mixture.

Chapter 14. Varying Densities for Some or All Crops in a Mixture

• Results in Chapter 5 of Volume I are extended to mixtures of three or more crops.

Chapter 15. Mixing Ability Effects When Individual Cultivar Responses are Available

• In Chapter 6 of Volume I, mixing ability effects were described and illustrated for two crops in the mixture. For three or more crops in a mixture, TDs and analyses are more complex and involved and require care in interpretation.

Chapter 16 . Intercrop Mixtures When Individual Crop Responses Are Not Available

• This chapter is an extension of the material given in Chapter 7 of Volume I. The TDs and analyses for three or more crops in a mixture are described and illustrated.

Chapter 17. Spatial and Density Arrangements

• In this chapter, we expand the material presented in Chapter 8 of Volume I. Density, spatial, and intimacy relations for three or more cultivars in a mixture become more complex than for two cultivars, as does the analysis.

Chapter 18. Some Analytical Variations for Intercropping Studies

• The extension of the material in Chapter 9 of Volume I requires extension of various indices from two cultivars to more than two. There is considerable literature on these methods for two crops, but the extension to three or more crops in a mixture has received little or no attention.

Chapter 19. Application of Intercropping Procedures to Other Fields

• The statistical designs and analyses developed herein are applicable in a wide variety of fields. In agriculture, rotation and sequential cropping studies can use the procedures developed in this book. In medicine involving mixtures of drugs (e.g., Waldholz, 1996), this is even discussed in the news media; many of the procedures can be applied directly and others only need minor modifications. Several other areas where the procedures described herein could be used are indicated.

Chapter 20. An Intercropping Bibliography

• Although this bibliography is extensive, like all bibliographies it is never complete. Many of the references listed could be obtained over the Internet, but it was deemed desirable to have a hard copy of the references available. The procedure for obtaining a copy of more than 3000 references is given in Chapter 20.

11.9 Literature Cited

Aina, P. O., R. Lal, and G. S. Tyler (1977). Soil and crop management in relation to soil erosion in rainforest of Western Nigeria. In *Soil Erosion Prediction and Control* (Editor: G. R. Foster), Special Publication 21, Soil and Water Conservation Society, Ankeny, IA, pp. 75-84.

Anonymous (1989). Decision reached on sustainable agriculture. *Agronomy News*, January, p. 15.

Balaam, L. N. (1986). Intercropping–past and present. In *Statistical Designs–Theory and Practice. Proceedings of a Conference in Honor of Walter T. Federer* (Editors: C. E. McCulloch, S. J. Schwager, G. Casella, and S. R. Searle), Biometrics Unit, Cornell University, Ithaca, NY, pp. 141-150.

de Wit, C. T. and J. P. van den Bergh (1965). Competition among herbage plants. *Netherlands J. Agric. Sci.* **13**, 212-221.

Ezumah, H. C. and W. T. Federer (1991). Intercropping cassava and grain legumes in humid Africa. 2. Cassava root yield, energy, monetary, and protein returns of system. BU-1115-M in the Technical Report Series of the Biometrics Unit, Cornell University, Ithaca, NY.

Ezumah, H. C. and N. R. Hullugalle (1989). Studies on cassava-based rotation systems in tropical Alfisol. *Agronomy Abstracts*, ASA, Madison, WI, pp. 53.

Federer, W. T. (1987). Statistical analysis for intercropping experiments. *Proceedings, Thirty-Second Conference on the Design of Experimnts in Army Research Development and Testing*, ARO-87-2, pp. 1-29.

Federer, W. T. (1989). Intercropping, developing countries, and tropical agriculture. *Biometrics Bull.* **6**, 22-24.

Federer, W. T. (1993a). Statistical design and analysis of intercropping experiments. In *Crop Improvement for Sustainable Agriculture* (Editors: M. B.

Callaway and C. A. Francis), University of Nebraska Press, Lincoln, pp. 234-249.

Federer, W. T. (1993b). *Statistical Design and Analysis for Intercropping Experiments. Volume I : Two Crops*. Springer-Verlag, New York.

Federer, W. T. and B. R. Murty (1987). Uses, limitations, and requirements of multivariate analyses for intercropping experiments. In *Biostatistics* (Editors: I. B. MacNeill and G. J. Umphrey), D. Reidl, Dortrecht, The Netherlands, pp. 269-283.

Federer, W. T. and S. J. Schwager (1982). On the distribution of land equivalent ratios. BU-777-M in the Technical Report Series, Biometrics Unit, Cornell University, Ithaca, NY.

Hullugalle, N. R. and H. C. Ezumah (1989). Effect of cassava-based cropping system and rotation on soil physical properties of an Alfisol in Southwestern Nigeria. *Proc. 6th International Conservation Conference*, Addis Ababa, Ethiopia.

Kass, D. C. L. (1978). Polyculture systems: review and analysis. *Cornell International Agriculture Bulletin 32*, CALS, Cornell University, Ithaca, NY, pp. 1-69.

Lal, R. (1989). Conservation tillage for sustainable agriculture: tropics vs. temperate environments. *Adv. Agric.* **42**, 85-195.

Mead, R. and J. Riley (1981). A review of statistical ideas relevant to intercropping research (with discussion). *J. Roy. Statist. Soc., Series A* **144**, 462-509.

Morales, M. F. M. (1993). Estimation and design of yield density models for intercropping experiments. Ph.D. Thesis, University of California at Riverside.

Pearce, S. C. and B. Gilliver (1978). The statistical analysis of data from intercropping experiments. *J. Agric. Sci., Cambridge* **91**, 625-632.

Pearce, S. C. and B. Gilliver (1979). Graphical assessment of intercropping methods. *J. Agric. Sci., Cambridge* **93**, 51-58.

Pimentel, D., G. Roderigues, T. Wang, R. Abrams, K. Goldberg, H. Staecker, E. Ma, L. Brueckner, L. Trovato, C. Chow, U. Govindarajulu, and S. Broerke (1994). Renewable energy: economic and environmental issues. *BioScience* **44**, 536-547.

The Holy Bible (1952). Revised Standard Edition, Nelson, Toronto.

Waldholz, M. (1996). Three-drug therapy may suppress HIV. *The Wall Street Journal, Medicine*, January 30.

Weil, R. R. (1990). Defining and using the concept of sustainable agriculture. *J. Agron. Ed.* **19**, 126-130.

Willey, R. W. and D. S. O. Osiru (1972). Studies on mixtures of maize and beans (*Phaseolus vulgaris*) with particular reference to plant populations. *J. Agric. Sci., Cambridge* **79**, 517-529.

CHAPTER 12

One Main Crop Grown with More Than One Supplementary Crop

12.1 Introduction

In Chapter 2 of Volume I, we discussed the situation for which there was *one* main crop grown with *one* supplementary crop. In this chapter, we consider the situation where 1, 2, . . . , v cultivars of a secondary crop or crops are grown with one main crop. For example, consider one line or variety of sugarcane which does not "close in," that is, form a canopy of shade, for 4 months after planting; when the sugarcane plants are small, within the first 4 months after planting, the plants do not fully utilize all the available space, water, and nutrients. In order to utilize this material more fully, short-season annuals are planted between the rows of sugarcane. Such crops as onions, cowpeas, beans, radishes, potatoes, and melons, alone or in combinations, have been used successfully with sugarcane. The supplementary crops must be such that the yield of the main crop is either relatively unaffected or is enhanced. Short-season crops may be grown simultaneously or in sequence during the first months of the sugarcane crop.

Another example where short-season annuals may be grown with a main crop is cassava (manioc, yucca). Since cassava plants start off slowly and the plants are relatively far apart, the land is not fully utilized during the first few months after the cassava has been planted. Greens, melons, cowpeas, beans, potatoes, etc., alone or in mixtures, have been used successfully as supplementary intercrops with the main crop cassava. Another example is using a grain crop, e.g., oats, in a grass–legume mixture. The grain crop has been called a "nurse crop," while the grass is included with the main crop legume to have a grass–legume hay. When paddy rice is the main crop, the edges around the paddy have been used to grow a variety of crops, including mixtures of crops. When rubber trees were the main crop, beans, cotton, and maize, alone and in mixtures, have been grown during the first year or two while the rubber trees in the plantation were being established. Legume–grass mixtures are grown as secondary crops in various types of fruit orchards.

In many situations, the correct choice and density of the supplementary crops will leave the main crop yield relatively unaffected. The benefit then would be the value of the supplementary crops, as no extra land is utilized. It is possible and not infrequent that the main crop yields may be increased by the presence of the supplementary crops. For example, in Nigeria, when cassava is intercropped with melons, its yield is actually increased. The reason is that the melons prevent erosion over and above that found in the sole crop cassava. The erosion-control aspects of melons more than offset any competition between melons and cassava for space, water, and nutrients. Thus, intercropping cassava with melons not only produces a partial crop of melons but it actually increases the yield of the main crop cassava. With long-season crops like sugarcane and cassava, legumes with nitrogen-fixation qualities forming nodules on the roots should be successful in enhancing the yields of cassava and sugarcane. These main crops would be able to utilize the nitrogen nodules left in the soil as they decomposed. The fertilizer replacement qualities of the legume may be quite beneficial for crops of this nature.

For several of the above situations, two or more supplementary crops may be grown in sequence rather than simultaneously. For example, one might have a sequence of melons–cowpeas, cowpeas–melons, maize–beans, etc. with cassava. With sugarcane, peas might be planted, and shortly after the pea plants are ready to flower, onions could be planted in between the rows of peas. Alternatively, the peas might be planted and harvested and then the onions would be planted. Many such schemes can be, and are, utilized depending on the crops, varieties, and environmental conditions.

When two or more supplementary crops are grown with a main crop, the number of possible combinations becomes large and much faster than the number of secondary crops. Response model equations also become much more complex, and the number of parameters increases over that when there is only one supplementary crop. Statistical design aspects for this situation are discussed in the following section. Some simple response model equations, estimators for parameters, and variance of estimates are discussed in Section 12.3 for $2, 3, \ldots, v$ supplementary crops grown with one main crop. It is shown how to add additional parameters to the response model equations; these may be of interest in certain situations. To illustrate the analyses of Section 12.3, we consider a barley experiment where the main crop, barley, is intercropped with one, with three, and with all six cultivars, as well as being grown as a sole crop. This is discussed in Section 12.4. In Section 12.5, we consider the case of c lines or varieties of one main crop, where each line of the main crop is intercropped with two or more supplementary crops. Two different experiment designs are described. Some comments on the example are given in Section 12.6 and some problems are given in Section 12.7.

12.2 Some Statistical Design Considerations

One possible treatment design for one main crop cultivar and v supplementary crops follows:

main crop grown as a sole crop,
main crop grown with each one of the v supplementary crops,
main crop grown with each of the $v(v-1)/2$ pairs of supplementary crops,
main crop grown with each of the $v(v-1)(v-2)/6$ triples of supplementary crops,
$\vdots$
main crop grown with each set of $v-1$ of the supplementary crops, and
main crop grown with all v supplementary crops.

The total number of treatments would be $\sum_{k=0}^{v}\binom{v}{k} = 2^v = N$. The experimenter would usually eliminate certain values of k and/or other combinations to reduce N considerably and would often know that certain combinations were undesirable or would not be used in practice. These usually would not be included in the experiment. There are many possible subsets of N and the experimenter should determine which subset to use to meet the goals of the experiment. For example, the treatment design could be a sole main crop, the main crop with each of the v supplementary crops, and the main crop with all possible pairs of the v supplementary crops for a total of $1+v+v(v-1)/2$ treatments. Results from fractional replication may be of use here (see, e.g., Cochran and Cox, 1957, Federer, 1967, Raktoe *et al.*, 1981). As an example, interest could center on only mixtures of size four and only main effects and two-factor interactions. Then, only $v(v-1)/2$ mixtures of the total number of combinations of $v(v-1)(v-2)(v-3)/24$ would be used. For $v = 8$, the fraction would be 28 out of 70 possible mixtures. Examples of fractional replicates appear in Chapters 15 and 16.

Also, it is possible that it would be desirable to replicate the sole crop treatment more frequently than the others. If all comparisons are to be made with the sole crop, then for r replications of each of the other treatments, the number of replications for the sole crop could be $\sqrt{N}$ to the nearest integer in each of r blocks in order to optimize variance considerations. If the experiment design were an incomplete block, the $\sqrt{N}$ sole crop experimental units would be scattered over the incomplete blocks, such that sole crop (experimental units) would appear m or $m+1$ times in an incomplete block, $m = 0, 1, 2, \ldots$. Such an arrangement would decrease the variance between sole crop and other combinations.

The relatively large number N of treatments possible should make the experimenter consider each treatment carefully for inclusion or exclusion in the experiment. Combinations that could reduce main crop yields below an acceptable level should be excluded from the experiment. Combinations which would not be used in practice should usually be excluded. Response model equations (12.1) to (12.5) do not depend on having all possible combinations for various values of $k = 0, 1, 2, \ldots, v$ or even for a particular value of k. If one uses equations (12.21) and (12.22) to add more parameters in (12.1) to (12.5), then all combinations for a particular value of k would be required.

Selection of treatments, densities of supplementary crops, and spatial arrangements require considerable thought. For example, suppose that maize and beans are the supplemental crops, and that paddy rice is the main crop. Further, suppose that the paddies are in a rectangular arrangement. On the paddy edges, we could consider some such arrangement as given in Figure 12.1. In this figure, m = maize alone, b = beans alone, and mb = maize and beans intercropped. What is the best way to fill in the unmarked edges to consider such treatments as $3b : m$, $2b : 2m$, $b : 3m$, $3b : mb$, $2b : 2mb$, $b : 3mb$? If interest centered only on m and b grown alone (i.e., not mb) from the paddy edges, use may be made of the Veevers and Zaraf (1982) approach or of other approaches described in Chapter 8 of Volume I. A balanced arrangement of m and b would be used such that the rice paddy had m on 0, 1, 2, 3, and 4 edges of the paddy and b on the remaining edges. In addition to whatever combinations were used above, sole plots of paddy rice, that is, nothing planted on the paddy edges, may be included.

If it were desired to vary the density of supplementary crops, as a preliminary step, the density of the supplementary crop along the edge of the paddy from zero to dense could be varied. Then, the rice yields of the row(s) adjacent to the

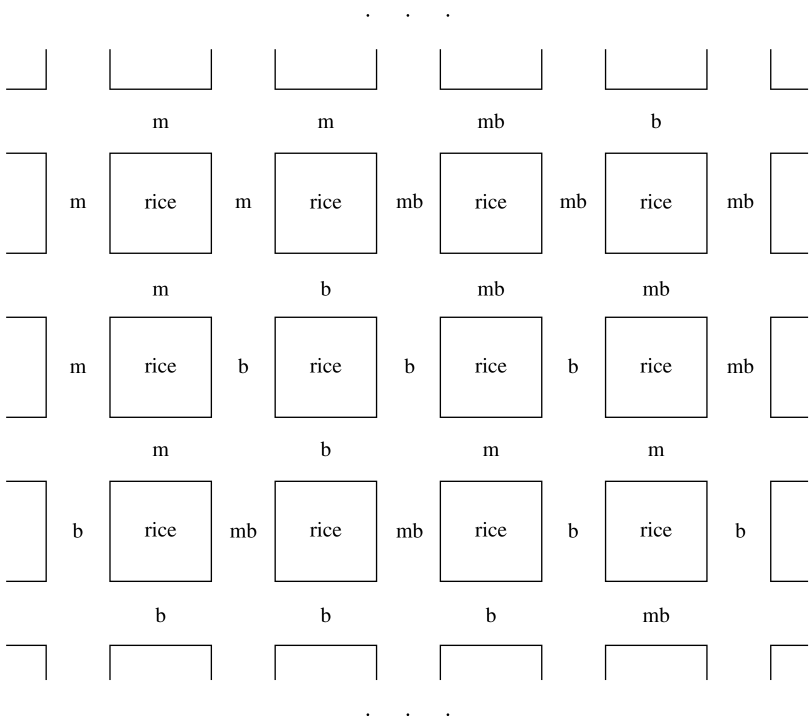

FIGURE 12.1. A possible arrangement of paddy rice with m (maize) and b (beans) planted on the paddy edges.

paddy edge could be collected. A plot of the yield of the rice against density of the supplemental crop along the paddy edge would indicate the highest density which did not affect rice yield. Such a scheme, using the parsimonious designs in Chapter 10 of Volume I, would require a minimum of experimental material, say r rice paddy edges. Then, a more detailed experiment could be conducted to determine the highest supplemental crop density not affecting main crop yields for a selected set of densities.

Some aspects of experiment design are discussed in Section 12.5. Each experiment should be considered on its own merit and should not be copied from another experiment. The axiom in Chapter 10 of Volume I, that the design should be tailored for the experiment rather than making the experiment fit a known or tabled design, applies here as well as in other intercropping experiments.

12.3 Response Model Equations

Consider an experiment involving a single cultivar, e.g., a given barley variety, for which the experiment design is a randomized complete blocks design (RCBD) with r blocks; the treatment design consists of the sole crop of the cultivar and mixtures of k of v additional lines, cultivars, or crop species with $k = 1, 2, \ldots, v$. The simplest possible response equations for the main crop, e.g., barley, would appear to be of the following form $h = 1, \cdots, r$:

<u>Sole crop cultivar</u>

$$Y_{h0} = \mu + \tau + \rho_h + \epsilon_h \tag{12.1}$$

<u>Cultivar plus one additional line (i)</u>

$$Y_{hi1} = \mu + \tau + \rho_h + \delta_i + \epsilon_{hi} \tag{12.2}$$

<u>Cultivar plus two additional lines (i and j)</u>

$$Y_{hij2} = \mu + \tau + \rho_h + \frac{1}{2}(\delta_i + \delta_j) + \gamma_{ij} + \epsilon_{hij} \tag{12.3}$$

<u>Cultivar plus three additional lines (i, j, and g)</u>

$$Y_{hijg3} = \mu + \tau + \rho_h + \frac{1}{3}(\delta_i + \delta_j + \delta_g) + \lambda_{ijg} + \epsilon_{hijg} \tag{12.4}$$

<u>Cultivar with all v additional lines</u>

$$Y_{hij\cdots v} = \mu + \tau + \rho_h + \delta_. + \pi_{12\cdots v} + \epsilon_{hij\cdots}. \tag{12.5}$$

$\mu + \tau$ is the mean of the barley cultivar (main crop) when grown as a sole crop; ρ_h is the hth replicate effect for the RCBD; δ_i is the effect on the barley cultivar when grown in a mixture with supplementary crop line i; γ_{ij} is the bi-specific mixing effect of lines i and j on the response for the barley cultivar; γ_{ijg} is the tri-specific mixing effect of lines i, j, and g on the response for the barley cultivar; $\delta_.$ is the average of the δ_i; $\gamma_{..}$ is the average of the γ_{ij}; $\lambda_{...}$ is the average of the λ_{ijg}; and so forth on down to $\pi_{12\cdots v}$, which is the v-specific mixing effect on the response for the barley cultivar; and the $\epsilon_h, \epsilon_{hi}, \epsilon_{hij}, \epsilon_{hijg}, \ldots, \epsilon_{hij\ldots}$ are random error deviations which are independently distributed normal variates with mean zero and variance σ_ϵ^2. Note that, in this form, only one additional effect is added as a line is added to the mixture and this effect is a mixture of several effects. For example, in the response equation for a mixture of three lines, the γ_{ijg} is composed of interactions between i and j, i and g, j and g, and among i, j, and g in mixtures of size three, as well as any difference in the δ_i from mixtures of one to mixtures of three. Also note that $\gamma_{..}$, $\gamma_{...}$, etc. are included in $\pi_{12\cdots v}$ but need not be if all combinations are present. Note the coefficients of $\frac{1}{2}$ and $\frac{1}{3}$ in (12.3) and (12.4), respectively. These coefficients are necessary to make the δ_i from experimental units with two and three supplementary cultivars comparable to the δ_i from experimental units with one supplementary cultivar. The same rationale holds for response equations when four, five, etc. supplementary crops appear in an experimental unit. The above holds when cultivars are allocated equal space in an experimental unit. The coefficients would need to be adjusted if unequal space were allocated to supplementary crops and would be proportional to the space occupied. Least squares solutions for the parameters in the response model equations are

$$\mu + \tau = \bar{y}_{\cdot 0} = \text{mean of sole crop yields,} \tag{12.6}$$

$$\hat{\delta}_i = \bar{y}_{\cdot i1} - \bar{y}_{\cdot 0}, \tag{12.7}$$

$$\hat{\delta}_{\cdot} = \sum_{i=1}^{v} \hat{\delta}_i / v = \bar{y}_{\cdot\cdot 1} - \bar{y}_{\cdot 0}, \tag{12.8}$$

$$\hat{\gamma}_{ij} = \bar{y}_{\cdot ij2} - (\bar{y}_{\cdot i1} + \bar{y}_{\cdot j1})/2, \tag{12.9}$$

$$\begin{aligned} \hat{\gamma}_{\cdot\cdot} &= 2 \sum_{i<j}\sum \hat{\gamma}_{ij} / v(v-1) \\ &= \bar{y}_{\cdots 2} - \bar{y}_{\cdot\cdot 1}, \end{aligned} \tag{12.10}$$

$$\hat{\lambda}_{ijg} = \bar{y}_{\cdot ijg3} - (\bar{y}_{\cdot i1} + \bar{y}_{\cdot j1} + \bar{y}_{\cdot g1})/3, \tag{12.11}$$

$$\begin{aligned} \hat{\lambda}_{\cdots} &= \frac{6}{v(v-1)(v-2)} \sum_{i<j<g}\sum\sum \hat{\lambda}_{ijg}, \\ &= \bar{y}_{\cdots\cdot 3} - \bar{y}_{\cdot\cdot 1}, \end{aligned} \tag{12.12}$$

$$\vdots$$

$$\hat{\pi}_{12\cdots v} = \bar{y}_{\cdot 12\cdots v} - \bar{y}_{\cdot\cdot 1}. \tag{12.13}$$

The standard dot notation is used to denote which subscript has been summed over, e.g., $\bar{y}_{\cdot i1}$ and $Y_{\cdot i1}$ are the mean and total, respectively, for the yield of the main crop when grown with crop i. The variances for the above solutions, under the assumption of homoscedasticity, are

$$V(\hat{\delta}_i) = \frac{2\sigma_\epsilon^2}{r}, \tag{12.14}$$

$$V(\hat{\delta}.) = \sigma_\epsilon^2 \left\{ \frac{1}{rv} + \frac{1}{r} \right\} = \frac{\sigma_\epsilon^2 (v+1)}{rv}, \tag{12.15}$$

$$V(\hat{\gamma}_{ij}) = 3\sigma_\epsilon^2 / 2r, \tag{12.16}$$

$$V(\hat{\gamma}..) = \sigma_\epsilon^2 \left\{ \frac{2}{rv(v-1)} + \frac{1}{rv} \right\} = \frac{\sigma_\epsilon^2 (v+1)}{rv(v-1)}, \tag{12.17}$$

$$V(\hat{\lambda}_{ijg}) = 4\sigma_\epsilon^2 / 3r, \tag{12.18}$$

$$V(\hat{\lambda}...) = \sigma_\epsilon^2 \left\{ \frac{6}{rv(v-1)(v-2)} + \frac{1}{rv} \right\}, \tag{12.19}$$

$$\vdots$$

$$V(\hat{\pi}_{12\cdots v}) = \frac{\sigma_\epsilon^2 (v+1)}{rv}. \tag{12.20}$$

Putting the solutions in terms of arithmetic means simplifies the determination of the variances for the various quantities in equations (12.6) to (12.13).

As stated previously, this appears to be about the simplest possible set of response equations when interactions are present. Additional parameters may be added easily to the response model equations. If all possible $v(v-1)/2$ pairs are present, the γ_{ij} parameters could be rewritten to obtain

$$\gamma_{ij} = \delta_{2i}^* + \delta_{2j}^* + \gamma_{2ij}^*, \tag{12.21}$$

where δ_{2i}^* is a general mixing effect of line i when pairs of lines are added to the main crop, and γ_{2ij}^* is a bi-specific mixing effect for pair ij. Likewise, the λ_{ijg} parameters may be changed to

$$\lambda_{ijg} = \delta_{3i}^* + \delta_{3j}^* + \delta_{3g}^* + \gamma_{3ij}^* + \gamma_{3ig}^* + \gamma_{3jg}^* + \lambda_{3ijg}^*, \tag{12.22}$$

where δ_{3i}^* is the general mixing effect of line i in mixtures of three lines, γ_{3ij}^* is a bi-specific mixing effect of lines i and j in the mixture ijg, γ_{3ig}^* and γ_{3jg}^* are defined in a manner similar to γ_{3ij}^*, and λ_{3ijg}^* is a tri-specific mixing effect in the mixture of lines i, j, and g. Solutions for each of the parameters in (12.21) and (12.22) may be obtained when all possible combinations are present and when $v > 3$ for (12.21) and $v > 5$ for (12.22). However, these should be of little interest in this chapter since we are only concerned with one main crop and how different mixtures affect the yield of that crop. They are of interest for the analyses in Chapters 15 and 16. It should be emphasized that (12.1) to (12.5) may be used when only a subset of all possible combinations are present.

For a given experiment in an RCBD with r blocks, wherein there is one sole crop, the main crop, and v supplementary crops, the treatment design could con-

TABLE 12.1. Analysis of Variance Table for RCBD with Response Model Equations (12.1) to (12.5). (The Treatment Degrees of Freedom Will Change If Other Than All Possible Combinations Are Present.)

Source of variation	df	Sum of squares
Total	rt	—
Correction for mean	1	Usual manner for a RCBD
Blocks	$r-1$	—
Treatments	$t-1$	—
Sole vs. rest	1	$\frac{Y_{.0}^2}{r} + \frac{(Y_{..1}+\cdots+Y_{....v})^2}{r(t-1)} - \frac{(Y_{.0}+Y_{..1}+\cdots+Y_{....v})^2}{rt}$
Among sole + singles	$v-1$	$\sum_{i=1}^{v} \frac{Y_{.i1}^2}{r} - \frac{Y_{..1}^2}{rv}$
Sole + Singles vs. rem	1	$\frac{Y_{..1}^2}{rv} + \frac{(Y_{...2}+\cdots+Y_{....v})^2}{r(t-v)-r} - \frac{(Y_{..1}+\cdots+Y_{....v})^2}{r(t-1)}$
Among sole + pairs	$\frac{v(v-1)}{2}-1$	$\sum\sum_{i<j} \frac{Y_{ij2}^2}{r} - \frac{2Y_{...2}^2}{rv(v-1)}$
Sole + pairs vs. remaining	1	$\frac{2Y_{...2}^2}{rv(v-1)} + \frac{2(Y_{....3}+\cdots+Y_{....v})^2}{r(2t-v^2-v-2)} - \frac{(Y_{...2}+\cdots+Y_{...v})^2}{r(t-v)-r}$
Among sole + triples	$\frac{v(v-1)(v-2)}{6}$	$\sum\sum\sum_{i<j} \frac{Y_{.ijg3}}{r} - \frac{6Y_{....3}^2}{rv(v-1)(v-2)}$
⋮		
Among sole + $v-1$	$v-1$	$\sum \frac{Y_{.if\cdots(v-1)}^2}{r} - \frac{(Y_{....(v-1)})^2}{rv}$
Sole + $(v-1)$ lines vs. all lines	1	$\frac{Y_{...(v-1)}^2}{rv} + \frac{Y_{....v}^2}{r} - \frac{(Y_{....(v-1)}+Y_{....})^2}{r(v+1)}$
Block × treatments	$(r-1)(t-1)$	By subtraction

sist of the one described in the first part of the previous section. An analysis of variance table for the above treatment design in an RCBD is given in Table 12.1. Response model equations (12.1) to (12.5) were assumed for this analysis of variance table. The block × treatment sums of squares may be partitioned into components of block × treatment contrasts to check for variance heterogeneity. Since the experimental units are of the same size, heterogeneity would not, in general, be suspected. However, it is always advisable to check for variance heterogeneity. The equations are directly extendible to incomplete block and row–column designs.

Instead of partitioning the treatment degrees of freedom as above and making F-tests, a multiple comparisons procedure could have been used for either significance testing or setting simultaneous confidence intervals as described in Chapters 2 and 3 of Volume I.

12.4 Example

S. Kaffka, Cornell University, conducted an experiment in large containers in a greenhouse at the University of Hohenheim, Stuttgart, West Germany, during March to July of 1980. A uniform stockpiled Filder clay–loam soil mixed with small amounts of peat moss and sand was used to fill the containers (boxes). All boxes were sown with sufficient barley seeds and seeds of the other six secondary species to establish a stand of 20 uniformly spaced barley plants and undersown plants according to the following pattern in one block of a randomized complete block design with $r = 3$ blocks (see Figure 12.2):

(i) 1 box with 20 barley plants and no secondary species,
(ii) 6 boxes in which 1 box contained 20 barley plants and 12 plants of 1 of the 6 species,
(iii) 20 boxes with 20 barley plants and 12 other plants which consisted of 4 plants (randomly allotted) from each of 3 of the 6 species and which was 1 of the 20 possible combinations of 6 species taken 3 at a time, and
(iv) 1 box which contained 20 barley plants and 2 plants of each of the 6 species.

All seeds were sown on one planting date, thinned to a single plant per position, and watered as necessary throughout the growing season. At the end of the growing season, 6 barley plants from the center of each box and all 12 plants of the secondary species were harvested and dry weights taken. A yield–density trial for barley and a replacement series of barley and lentils were also included in the experiment as a partial check on the model employed. The data for seed weight of the six barley

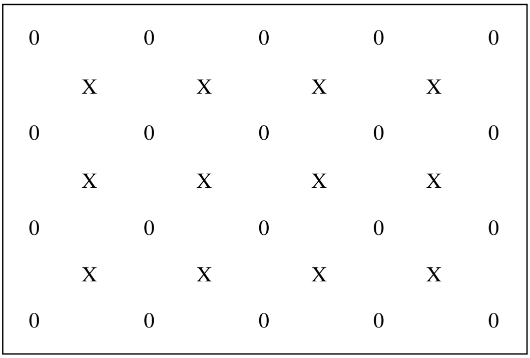

FIGURE 12.2. Experimental unit arrangement of 20 barley plants, denoted by 0, and 12 plants of 0, 1, or more of 6 cultivars denoted by X.

TABLE 12.2. Barley Grain Weight in Grams for Kaffka Experiment.

	Block (seed weight, g)				
Treatment	1	2	3	Total	Mean
Barley	22.7	16.8	18.9	58.4	19.47
Barley $+M_1$	19.1	12.9	18.5	50.5	18.83
Barley $+M_2$	19.2	15.3	18.2	52.7	17.57
Barley $+M_3$	17.4	16.9	20.4	54.7	18.23
Barley $+M_4$	12.8	23.1	26.2	62.1	20.70
Barley $+M_5$	15.6	25.7	15.2	56.5	18.83
Barley $+M_6$	18.6	16.7	18.6	53.9	17.97
Barley $+M_1 + M_2 + M_3$	15.1	22.1	23.7	60.9	20.30
Barley $+M_1 + M_2 + M_4$	16.6	22.0	18.8	57.4	19.13
Barley $+M_1 + M_2 + M_5$	19.8	22.8	23.6	66.2	22.07
Barley $+M_1 + M_2 + M_6$	19.6	19.2	19.2	58.0	19.33
Barley $+M_1 + M_3 + M_4$	17.0	23.0	25.2	65.2	21.73
Barley $+M_1 + M_3 + M_5$	17.2	15.1	21.5	53.8	17.93
Barley $+M_1 + M_3 + M_6$	18.3	21.1	19.5	58.9	19.63
Barley $+M_1 + M_4 + M_5$	16.0	20.9	20.8	57.7	19.23
Barley $+M_1 + M_4 + M_6$	21.3	24.6	22.3	68.2	22.73
Barley $+M_1 + M_5 + M_6$	14.2	17.5	20.1	51.8	17.27
Barley $+M_2 + M_3 + M_4$	24.0	23.2	18.6	65.8	21.93
Barley $+M_2 + M_3 + M_5$	15.7	15.4	19.3	50.4	16.80
Barley $+M_2 + M_3 + M_6$	16.2	20.2	26.2	62.6	20.87
Barley $+M_2 + M_4 + M_5$	19.3	17.7	22.7	59.7	19.90
Barley $+M_2 + M_4 + M_6$	21.3	29.4	24.6	73.3	25.10
Barley $+M_2 + M_5 + M_6$	22.8	16.9	22.0	61.7	20.57
Barley $+M_3 + M_4 + M_5$	16.9	26.0	16.1	57.0	19.67
Barley $+M_3 + M_4 + M_6$	21.6	15.0	21.3	57.9	19.30
Barley $+M_3 + M_5 + M_6$	23.9	19.2	21.1	64.2	21.40
Barley $+M_4 + M_5 + M_6$	18.9	24.0	18.2	61.1	20.37
Barley $+M_1 + M_2 + M_3$ $+M_4 + M_5 + M_6$	17.1	19.4	20.3	56.8	18.93
Total	518.2	562.1	581.1	1661.4	
Mean	18.51	20.08	20.75		19.78

plants are given in Table 12.2, but treatments described in the preceding sentence were omitted from the table. The 3 responses recorded were seed weight of 6 barley plants, plant dry weight of 6 barley plants, and total dry weight of 12 additional plants. In addition, mustard was sown soon after harvest in the same containers. It was grown as a check for any residual effects in soil nitrogen as a consequence of species mixtures. The mustard was harvested at flower stage, dried, and weighed.

The six secondary species were

1. *Avena fatua* (wild oat)
2. *Coriander sativa* (coriander)
3. *Lens esculentum* (lentils)

4. *Lotus corniculatus* (birdsfoot trefoil)
5. *Medicago sativa* (alfalfa)
6. *Matricaria chamomille* (chamomile)

The seed weight in grams of six barley plants is presented in Table 12.2. The treatment and block totals and means are also given. From these, one may compute residuals for a two-way array as

$$Y_{ij} - \bar{y}_{i\cdot} - \bar{y}_{\cdot j} + \bar{y}_{\cdot\cdot} = \hat{e}_{ij}$$

or

$$(rvY_{ij} - rY_{i\cdot} - vY_{\cdot j} + Y_{\cdot\cdot})/rv = \hat{e}_{ij}.$$

The first formula is subject to rounding errors, whereas the second form is not. The residuals times rv sum to zero exactly in any row or any column of the table using the second form. The frequency distribution of the 84 $\hat{e}_{ij}$'s is given in Figure 12.3. A rather symmetrical distribution, 43 negatives and 41 positives, was obtained with no unusual outliers, although the 3 residuals greater than 6 accounted for 21% of the total residual sum of squares. One could check on the relation between treatment means, $\bar{y}_{i\cdot}$, and sums of squares, $\sum_{j=1}^{r} \hat{e}_{ij}^2$, using Spearman's rank correlation. (D.S. Robson, Cornell University, and C.L. Wood, University of Kentucky, have shown that this follows Spearman's rank correlation.) First rank the means from 1 to 28; compute the 28 $\sum_{j=1}^{3} \hat{e}_{ij}^2$ in Table 12.3, and then rank them. Take the difference d_i in ranks. Then, Spearman's rank correlation is computed as

$$r_S = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)} = 1 - \frac{6\sum_{i=1}^{28} d_i^2}{28(28^2-1)} = 1 - \frac{6(3004)}{28(783)} = 0.18;$$

$r_S = 0.18$ is considerably smaller than $r_{.05}(26\text{ d.f.}) = 0.374$. Hence, the treatment means and variances are considered to be uncorrelated. In light of the above evidence, no transformation of seed weight was considered necessary to stabilize variances, which is required for F-tests.

The treatment means are presented in Table 12.2. The means of all mixtures involving a line, denoted as $M_{i\cdot\cdot}$, are presented in Table 12.4. All of these means are larger than the barley sole crop mean, 19.47. A graphical presentation depicting this is given in Figure 12.4. The chances of obtaining six out of six above the sole crop mean, given that the null hypothesis is true, is rather small. Since the expected number for the null hypothesis would be three, $\chi^2(1\text{ d.f.}) = (6-3)^2/3 + (0-3)^2/3 = 6 \doteq \chi^2_{.015}(1\text{ d.f.})$. Correcting this χ^2 for continuity gives $\chi^2(1\text{ d.f.}) = (2.5^2 + 2.5^2)/3 = 4.17 \doteq \chi^2_{.04}(1\text{ d.f.})$. The means for barley with a single line are also plotted on the graph in Figure 12.4. Here, we note that five out of the six are lower in yield than the barley sole crop. This gives $\chi^2(1\text{ d.f.}) = 8/3$, which, when corrected for continuity by reducing the deviation by $1/2$, becomes $1.4 \doteq \chi^2_{.20}(1\text{ d.f.})$. Also, in Table 12.4, means of pairs of cultivars are presented. Here, we note that 13 of the 15 pair means exceed the value 19.47 for the barley sole crop seed weight. Based on a null hypothesis of no effect, 7.5 would be the expected value. $\chi^2(1\text{ d.f.}) = [(13-7.5)^2 + (2-7.5)^2]/7.5 = 8.1$, which is a

TABLE 12.3. Residuals $(84Y_{hi} - 28Y_{\cdot i} - 3Y_{h\cdot} + Y_{\cdot\cdot})/84 = \hat{e}_{ij}$ for Data of Table 12.2

Treatment	Block 1	Block 2	Block 3
Barley	4.505	−2.963	−1.542
Barley $+M_1$	3.538	−4.230	0.692
Barley $+M_2$	2.905	−2.563	−0.342
Barley $+M_3$	0.438	−1.630	1.192
Barley $+M_4$	−6.629	2.104	4.525
Barley $+M_5$	−1.962	6.570	−4.608
Barley $+M_6$	1.905	−1.563	−0.342
Barley $+M_1 + M_2 + M_3$	−3.929	1.504	2.425
Barley $+M_1 + M_2 + M_4$	−1.262	2.570	−1.308
Barley $+M_1 + M_2 + M_5$	−0.995	0.437	0.558
Barley $+M_1 + M_2 + M_6$	1.538	−0.430	−1.108
Barley $+M_1 + M_3 + M_4$	−3.462	0.970	2.492
Barley $+M_1 + M_3 + M_5$	0.538	−3.130	2.592
Barley $+M_1 + M_3 + M_6$	−0.062	1.170	−1.108
Barley $+M_1 + M_4 + M_5$	−1.962	1.370	0.592
Barley $+M_1 + M_4 + M_6$	−0.162	1.570	−1.408
Barley $+M_1 + M_5 + M_6$	−1.795	−0.063	1.858
Barley $+M_2 + M_3 + M_4$	3.338	0.970	−4.308
Barley $+M_2 + M_3 + M_5$	0.171	−1.696	1.525
Barley $+M_2 + M_3 + M_6$	−3.395	−0.963	4.358
Barley $+M_2 + M_4 + M_5$	0.671	−2.496	1.825
Barley $+M_2 + M_4 + M_6$	−2.529	4.004	−1.475
Barley $+M_2 + M_5 + M_6$	3.505	−3.963	0.458
Barley $+M_3 + M_4 + M_5$	−1.495	6.037	−4.542
Barley $+M_3 + M_4 + M_6$	3.571	−4.596	1.025
Barley $+M_3 + M_5 + M_6$	3.771	−2.496	−1.275
Barley $+M_4 + M_5 + M_6$	−0.195	3.337	−3.142
Barley $+M_1 + M_2 + M_3 + M_4 + M_5 + M_6$	−0.562	0.170	0.392

relatively large value for χ^2. This is evidence that mixtures of barley with three of the six species produced higher yields of barley grain than did barley alone or with barley plus one of the six species.

To use the t (lsd) or range (hsd) statistics, one may compute an (where lsd = least significant difference (5%) and hsd is honestly significant difference (5%))

$$\begin{aligned} \text{lsd} &= \bar{y}_{\cdot i} - \bar{y}_{\cdot i'} \pm t_\alpha(54 \text{ d.f.})\sqrt{10.8804(1/3 + 1/3)}, \ \alpha = .05 \\ &= \bar{y}_{\cdot i} - \bar{y}_{\cdot i'} \pm 2(2.693) = \bar{y}_{\cdot i} - \bar{y}_{\cdot i'} \pm 5.39, \end{aligned}$$

and an

$$\begin{aligned} \text{hsd} &= \bar{y}_{\cdot i} - \bar{y}_{\cdot i'} \pm q_{\alpha,28,54} s_{\bar{y}_{\cdot i}}, \ \alpha = .05 \\ &= \bar{y}_{\cdot i} - \bar{y}_{\cdot i'} \pm 5.52\sqrt{10.8804/3} \end{aligned}$$

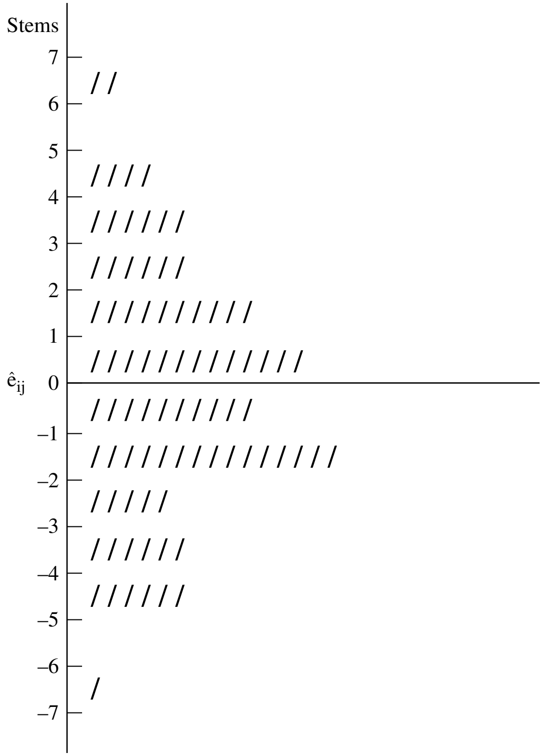

FIGURE 12.3. Frequency distribution of $\hat{e}_{ij}$.

$$= \bar{y}_{\cdot i} - \bar{y}_{\cdot i'} \pm 10.51.$$

The means of barley + one line versus the mean of barley + three lines may be compared using a t-statistic as follows:

$$t = (20.26 - 18.36)/\sqrt{10.8804(1/18 + 1/60)}$$
$$= 1.90/0.8865 = 2.14 > t_{.05}(54 \text{ d.f.}) = 2.00.$$

Sums of squares and F-statistics for a number of comparisons of the form described in Table 12.1 are given in Table 12.5. The only significant contrast was in the mean of mixtures of three lines with barley versus the mean of single lines with barley as given by the above t-statistic. Note that $\sqrt{F} = \sqrt{4.63} = 2.14 = t$ within rounding errors.

The coefficient of variation $\sqrt{10.8804}/19.78 = 16.7\%$ was rather high. It may be that some of the effects would be distinguishable on another scale. Taking logarithms, the coefficient of variation was 5.7%, but it was not ascertained if effects were more distinguishable, since the treatment sum of squares was not partitioned. This is left as an exercise for the reader. The ratio of the treatment

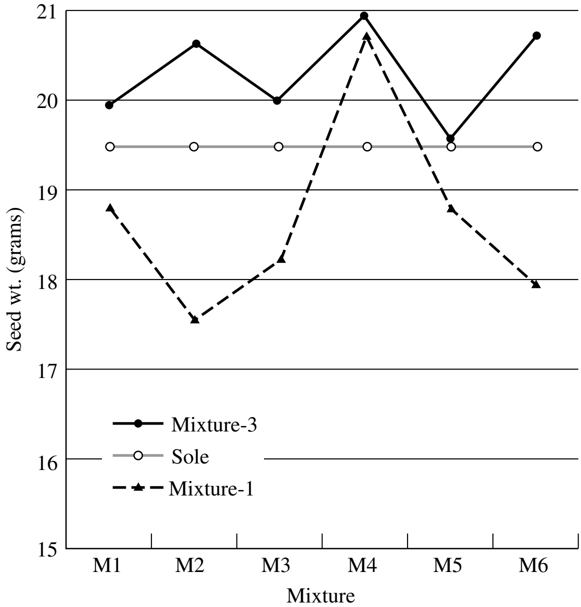

FIGURE 12.4. Comparison of sole crop barley with barley plus one other line and with means by line of barley planted with three of the six lines.

mean square to the block × treatment mean square was a little larger than unity, whereas it was a little smaller than one for the untransformed data.

To check for additivity, one may use Tukey's one-degree-of-freedom for the nonadditivity approach. Instead of the sum of products described in Chapter 10 of Volume I, we use another form which is simpler computationally, since we already have the residuals $\hat{e}_{ij}$ in Table 12.3. The form used for a one-degree-of-freedom sum of squares was

$$\left[\sum_i \sum_j \hat{e}_{ij} Y_{i\cdot} Y_{\cdot j}\right]^2 \Big/ rv \left(\frac{\Sigma Y_{i\cdot}^2}{r} - \frac{Y_{\cdot\cdot}^2}{rv}\right)\left(\frac{\Sigma Y_{\cdot j}^2}{r} - \frac{Y_{\cdot\cdot}^2}{rv}\right)$$

$$= \left[\sum_i \sum_j \hat{e}_{ij}(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})(\bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot})\right]^2 \Big/ \sum_i (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2 \sum_j (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2,$$

which, for the data in Tables 12.2 and 12.3, is equal to

$$3681.57^2/84(292.10)(74.34) = 7.43.$$

This mean square is less than the blocks × treatment mean square, 10.8804, indicating no evidence of departure from additivity. Since there appears to be no variance heterogeneity and no nonadditivity, it would appear that the responses Y_{ij} should not undergo any transformation for purposes of statistical analyses.

TABLE 12.4. Barley Seed Weight (g) by Individual Lines and Pairs of Lines in Combinations of Three Lines Plus Barley.

Total and mean of line over all combinations involving line i		
Line	Total	Mean, $M_{i\cdot\cdot}$
1	598.1	19.94
2	618.0	20.60
3	598.7	19.96
4	627.3	20.91
5	585.6	19.52
6	619.7	20.66
Total	3647.4	20.26

Totals and means of line pairs withh all others

Pair	Total	Mean	Pair	Total	Mean
12	242.5	20.21	26	257.6	21.47
13	238.8	19.90	34	247.9	20.66
14	248.5	20.71	35	227.4	18.95
15	229.5	19.12	36	243.6	20.30
16	236.9	19.74	45	237.5	19.79
23	239.7	19.98	46	262.5	21.88
24	258.2	21.52	56	238.8	19.90
25	238.0	19.83	Total	3647.4	

Analysis of variance for data of Table 12.2

Source of variation	Degrees of freedom	Sum of squares	Mean square
Total	84	33,814.10	
Correction for mean	1	32,860.12	
Blocks	2	74.34	37.17
Treatments	27	292.10	10.82
Block × treatment	54	587.54	10.8804
Barley vs. barley + a line	1	3.17	3.17
Barley vs. barley + a triple	1	1.81	1.81

This agrees with using Spearman's rank order correlation above to check for a relationship between rank of means and rank of residual sums of squares.

Solutions for values of the parameters given by equations (12.6) to (12.13) for the data of Table 12.2 are given in Table 12.6, along with variances. The computations for obtaining solutions are relatively simple, as they utilize only simple arithmetic means. The main crop cultivar effect is computed from (12.6) as

$$\bar{y}_{\cdot 0} = \hat{\mu} + \hat{\tau} = 58.4/3 = 19.47.$$

TABLE 12.5. Partitioning of Treatment Sum of Squares as Outlined in Table 12.1

Source of variation	Degrees of freedom	Sum of squares	Mean square	F
Treatments	27	292.10	10.82	0.99
Among single crops with barley = singles	5	26.49	5.30	0.49
Among triplets with barley = triplets	19	212.63	11.19	1.03
Barley sole vs. rest	1	0.30	0.30	0.03
Singles vs. rest including barley and all six	1	47.63	47.63	4.38
Triplets vs. all six	1	5.05	5.05	0.46
Triplets vs. singles	1	50.39	50.39	4.63
Sole crop vs. singles	1	3.17	3.17	0.29
Block × treatments	54		10.8804	

$F_{.10}(1, 54) = 2.80$	$F_{.25}(5, 54) = 1.38$	$F_{.25}(19, 54) = 1.26$
$F_{.05}(1, 54) = 4.02$	$F_{.05}(5, 54) = 2.38$	$F_{.05}(19, 54) = 1.76$
$F_{.01}(1, 54) = 7.12$	$F_{.01}(5, 54) = 3.37$	$F_{.01}(19, 54) = 2.24$

The estimated general mixing effect for an intercrop is obtained from (12.7). For cultivar 1, the computations are

$$\hat{\delta}_1 = \bar{y}_{\cdot 11} - \bar{y}_{\cdot 0} = 16.83 - 19.47 = -2.64.$$

The remaining cultivar general mixing effects are given in Table 12.6. From equation (12.11), we obtain solutions for the $\hat{\lambda}_{ijg}$'s; for example,

$$\hat{\lambda}_{123} = 20.30 - (16.83 + 17.57 + 18.23)/3 = 2.76.$$

The remaining $\hat{\lambda}_{ijg}$ values are given in Table 12.6. Note that 17 of 20 of these values are positive, the 3 negative values being near zero. Since such was the case and since all possible combinations of six cultivars taken three at a time were present, one should use equation (12.22) and obtain solutions for the δ^*_{3i}, the γ^*_{3ij}, and the λ^*_{3ijg} parameters. Such an occurrence most likely indicates that the δ_i and δ^*_{3i} parameters are different, whereas this fact was not taken into account in equations (12.1) to (12.13). This additional reparameterization will be done in Chapters 15 and 16.

12.5 Several Cultivars of Primary Interest

Suppose that c lines or cultivars are of primary interest and that v lines or cultivars of secondary interest are being considered. For example, suppose that c lines of barley, which will be grown with k of v supplementary cultivars, $k = 0, 1, 2, , \ldots, v$, are of interest. A split plot experiment design could be used in which

(i) the c lines or cultivars of primary interest form the whole plot or

TABLE 12.6. Solutions and Variances for Parameters in Response to Model Equations (12.1), (12.2), (12.4), and (12.5).

Equation	Solution	Equation	Variance	$\sqrt{\text{Var}}$
(12.6)	$\bar{y}_{.0} = 19.47$			
(12.7)	$\bar{\delta}_1 = -2.64$			
	$\bar{\delta}_2 = -1.90$			
	$\bar{\delta}_3 = -1.24$	(12.14)	$2(10.8804)/3 = 7.254$	2.69
	$\bar{\delta}_4 = 1.23$			
	$\bar{\delta}_5 = -0.64$			
	$\bar{\delta}_6 = -1.50$			
(12.8)	$\hat{\delta}_{.} = -1.115$	(12.15)	$7(10.8804)/18 = 4.231$	2.06
(12.11)	$\hat{\delta}_{123} = 2.76$			
	$\hat{\delta}_{124} = 0.77$			
	$\hat{\delta}_{125} = 4.32$			
	$\hat{\delta}_{126} = 1.88$			
	$\hat{\delta}_{134} = 3.14$			
	$\hat{\delta}_{135} = -0.03$			
	$\hat{\delta}_{136} = 1.96$			
	$\hat{\delta}_{145} = 0.44$			
	$\hat{\delta}_{146} = 4.23$			
	$\hat{\delta}_{156} = -0.61$	(12.18)	$4(10.8804)/9 = 4.836$	2.20
	$\hat{\delta}_{234} = 3.10$			
	$\hat{\delta}_{235} = -1.41$			
	$\hat{\delta}_{236} = 2.94$			
	$\hat{\delta}_{245} = 0.87$			
	$\hat{\delta}_{246} = 6.36$			
	$\hat{\delta}_{256} = 2.44$			
	$\hat{\delta}_{345} = 0.41$			
	$\hat{\delta}_{346} = 0.33$			
	$\hat{\delta}_{356} = 3.06$			
	$\hat{\delta}_{456} = 1.20$			
(12.12)	$\hat{\delta}_{...} = 1.908$	(12.19)	$13(10.8804)/180 = 0.786$	0.89
(12.13)	$\hat{\pi}_{123456} = 0.58$	(12.20)	$7(10.8804)/18 = 4.231$	2.06

(ii) the v supplementary crops form the whole plots.

The choice would depend on contrasts of primary interest. If a mixture combination for each cultivar of primary interest was desired, then use (i). If, on the other hand, it was desired to have more information on the c cultivars of primary interest, then use (ii). If all contrasts were of equal interest, then a complete block or an incomplete block design would be indicated.

Analyses of variance for situations (i) and (ii) above are given in Tables 12.7 and 12.8. It is recommended that analyses of variance of the form of Table 12.1 be performed for each line of the main crop prior to combining results as in Table 12.7. For (ii), analyses of variance should be obtained for each whole plot treatment

TABLE 12.7. Analysis of Variance for Lines or Cultivars of Main Crops as Whole Plots Arranged in an RCBD.

Source of variation	Degrees of freedom	Description
Total	rct	
Correction for mean	1	
Blocks $= R$	$r-1$	
Whole plot treatments, cultivars of secondary interest $= W$	$c-1$	
Blocks $\times$ whole plots	$(r-1)(c-1)$	Error line for whole plots
Treatments $= T$ (primary interest)	$t-1$	
T_1 Sole vs. rest	1	
T_2 Among sole + singles $= S$	$v-1$	
T_3 S vs. remaining treatments	1	
T_4 Among sole + pairs $= P$	$[v(v-1)-2]/2$	
T_5 P vs. remaining treatments	1	
T_6 Among sole + triplets	$[v(v-1)(v-2)-6]/6$	
$\vdots$	$\vdots$	
T_{n-1} Among sole $+v-1$ of v crops	$v-1$	
$T_n(v-1)$ crops + sole vs. v crops + sole	1	
$T \times W$	$(c-1)(t-1)$	
$W \times T_1$	$(c-1)$	
$W \times T_2$	$(c-1)(v-1)$	
$W \times T_3$	$c-1$	
$W \times T_4$	$(c-1)(v^2-v-2)/2$	
$\vdots$	$\vdots$	
$W \times T_n$	$c-1$	
$T \times R : W$	$c(r-1)(t-1)$	Error line for split plots

before combining the results for all whole plots. Standard statistical software for obtaining analyses for split plot designs may be used for these analyses. To obtain some of the sums of squares, a contrast statement is needed.

12.6 Some Comments

In the previous chapter, it was stated that one should not generalize from mixtures of two to mixtures of four, that one could expect surprises when analyzing data from a mixture experiment, and that one should not generalize from cultivar to cultivar. The example discussed in this chapter bears out these comments. When

TABLE 12.8. Analysis of Variance for a Supplementary Crop or Mixture of Supplementary Crops as the Whole Plots and with Lines of Cultivars of the Lines as the Split Plots.

Source of variation	Degrees of freedom	Description
Total	rct	
Correction for mean	1	
Blocks = M	$r - 1$	
Whole plot treatments, supplementary of secondary interest = T	$(t - 1)$	
$T \times R$	$(r - 1)(v - 1)$	Error line for whole plots
Cultivars of primary interest = W	$c - 1$	
$W \times T$	$(v - 1)(c - 1)$	
$M \times T : T$	$t(r - 1)(c - 1)$	Error line for split plots

Note that the partition of the degrees of freedom for T in Table 12.7 could be used for T in the above table. Likewise, the $T \times W$ interaction could also be partitioned as described in Table 12.7.

this particular barley variety was grown with one of the particular six cultivars, the yield was decreased for five of the six, relative to sole crop yield. The reverse was true for the barley variety grown with 3 of the 6 cultivars where 13 of the 20 mixtures of 4 outyielded the sole crop. Also, when averages of all mixtures of four in which one of the six cultivars was obtained, all six were above the sole crop average, 19.47 (see Table 12.4 and Figure 12.4). If a prediction had been made from mixtures of two for mixtures of four, it would have been predicted that mixtures of four would decrease yields. An error would have been made.

Such a result as discussed above for mixtures of two versus mixtures of four crops came as a surprise. Another surprise was that when the mixture contained barley plus all six cultivars, the barley yields were below the sole crop mean, i.e., 18.93 vs. 19.47. If these results are repeatable, they are interesting biological phenomena concerning species competition and ecology. Another surprise was that the 12 extra plants did not always decrease the yield of barley as this author would have presumed. The 12 extra plants should have exerted considerable stress on the barley plants, but this did not always materialize.

Even if these results are repeatable when the experiment is repeated, it would not be correct to generalize to other barley varieties and to other cultivars. The results are specific for this particular barley variety and the particular collection of the six supplementary crops used in the experiment. It is possible that the results are more general than indicated, but experiments should be conducted to confirm this.

TABLE 12.9.

Treatment	Block (plant weight in grams)		
	1	2	3
Barley	43.90	33.09	40.98
Barley $+M_1$	38.16	27.48	38.28
Barley $+M_2$	38.35	35.66	36.23
Barley $+M_3$	33.59	33.78	40.10
Barley $+M_4$	30.46	49.47	53.49
Barley $+M_5$	31.39	55.69	31.38
Barley $+M_6$	41.19	37.07	36.69
Barley $+M_1 + M_2 + M_3$	33.99	46.22	46.03
Barley $+M_1 + M_2 + M_4$	36.04	46.61	37.60
Barley $+M_1 + M_2 + M_5$	37.97	51.60	49.28
Barley $+M_1 + M_2 + M_6$	42.61	41.73	40.59
Barley $+M_1 + M_3 + M_4$	34.12	42.43	56.54
Barley $+M_1 + M_3 + M_5$	34.40	32.23	40.33
Barley $+M_1 + M_3 + M_6$	38.32	39.69	38.27
Barley $+M_1 + M_4 + M_5$	32.77	46.08	47.89
Barley $+M_1 + M_4 + M_6$	42.50	47.24	48.98
Barley $+M_1 + M_5 + M_6$	33.31	39.68	43.37
Barley $+M_2 + M_3 + M_4$	47.67	48.18	36.69
Barley $+M_2 + M_3 + M_5$	31.78	34.39	37.72
Barley $+M_2 + M_3 + M_6$	34.19	45.61	55.18
Barley $+M_2 + M_4 + M_5$	41.88	41.21	49.18
Barley $+M_2 + M_4 + M_6$	40.69	60.72	46.91
Barley $+M_2 + M_5 + M_6$	43.17	36.70	44.69
Barley $+M_3 + M_4 + M_5$	34.39	54.95	34.61
Barley $+M_3 + M_4 + M_6$	43.84	34.55	48.75
Barley $+M_3 + M_5 + M_6$	46.68	43.37	42.23
Barley $+M_4 + M_5 + M_6$	39.36	48.37	36.23
Barley $+M_1 + M_2 + M_3 + M_4 + M_5 + M_6$	34.73	40.03	39.78

12.7 Problems

12.1 Use the data of Table 12.2 for barley as a sole crop and with barley plus each of the six cultivars to obtain $3 \times 7 = 21$ observations. Conduct the analyses given in this chapter and interpret the results of your analyses. Prepare the necessary figures to aid with your interpretation.

12.2 Apply a logarithmic transformation to the data in Table 12.2 and conduct all the analyses described in this chapter. From these analyses, are the reasons for conducting analyses on untransformed data substantial? Why or why not?

12.3 Conduct the analyses described in this chapter on the following data. Use appropriate graphs and figures to aid in your interpretation of the data. How well do the results of these analyses agree with those on seed weight in Table 12.2?

12.8 Literature Cited

Cochran, W.G. and G.M. Cox (1957). *Experimental Designs*, 2nd ed. John Wiley and Sons, Inc., New York.

Federer, W.T. (1967). *Experimental Design – Theory and Application*. Oxford & IBH Publishing Company, Calcutta. (Reprint of book published by Macmillan, New York, 1955).

Raktoe, B.L., A. Hedayat, and W.T. Federer (1981). *Factorial Designs*. John Wiley and Sons, Inc., New York.

Veevers, A. and M. Zaraf-Yab (1982). A two-variety square lattice design balanced under Besag's coding scheme. *J. Roy. Statist. Soc., B* **44**, 47-48.

CHAPTER 13

Three or More Main Crops—Density Constant

13.1 Introduction

In many situations involving intercropping, three or more of the crops in a mixture may be considered to be the main crops. The grower is interested in a farming system and not necessarily in how each crop in the mixture performs. A desirable system would be one yielding the highest return in calories, in protein, in land use, in crop value (monetary or otherwise), and/or in some other evaluation of the system. From this point of view, all crops in a mixture would be considered to be main crops. Considering crops to be main crops need not imply that they are equal in value but that the grower will use these crops in a farming system.

There are many types of systems, as is partially demonstrated by the five examples given in the following sections. Great variation in systems exists. The experimenter should always ascertain which set of response model equations and which statistical analyses are appropriate to meet the type and goals of the particular experiment involved.

In Section 13.2, some comments on treatment design are given and illustrated with four examples. Treatment designs are different for the four examples and even more so for Example 13.5. Response model equations for each crop are given in Section 13.3. Estimators for the various parameters are presented along with an analysis of variance. The results are applied to a set of data from a mixture experiment. These analyses are for the yields of the individual crops in the spirit of the previous chapter.

Since a grower would be interested in a system, methods of combining the crop responses are given in Section 13.4. These results are applied to the data from Examples 13.3 and 13.5 and for a set of data for all possible combinations of three crops. Land equivalent ratios are generalized from two to v crops. Also, other created variables such as total calories, total protein, and total value are given for v crops. This is a generalization of the results presented in Chapter 4 of Volume I. Rather than use actual conversion factors, a ratio of coefficients is used which largely eliminates year-to-year variation in variables such as price. This requires selecting one of the crops as a base crop and the created variables will then be relative land equivalent ratios, relative values, etc. For comparative purposes, these relative variables are appropriate, and the ratios of yields, prices, etc. are considerably less variable than are actual values.

Some comments on the results from the experiments are given in Section 13.5. Some results are expected and others not. The last section is a derivation of some of the results in Section 13.3 and was relegated to an appendix rather than including it in the text.

13.2 Treatment Design

The treatment design given in Chapter 12, or subsets of it, may be used in this chapter as well. However, there are many variations that are used in intercropping experiments. Four other examples are described below. These have been reported by Aidar (1978) in his thesis and were made available through the courtesy of J. G. de Silva, EMBRAPA, in 1980.

Example 13.1. The treatment design consisted of the following eight treatments:

A cotton grown in sole crop
B cotton + 2 rows of maize
C cotton + 2 rows of beans
D cotton + 1 row of maize + 1 row of beans
E cotton + 2 rows of maize + 2 rows of beans
F cotton + 1 row of maize
G cotton + 1 row of beans
H cotton + 1 row of maize + 1 row of beans

Treatment H was different from D in that the bean plants were planted in with the maize plants, whereas, in D, there was one row of maize and one row of beans. This example could have been used in Chapter 12 if cotton was the main crop, say, and maize and beans, say, were the supplementary crops.

Example 13.2. The following treatment design and yields were obtained from an intercropping experiment involving castrol beans, maize, and beans. The data come from an experiment grown in Ibititá, Bahia, Brazil, in 1972.

	Treatment	Yield (kg/ha) Castrol bean	Maize	Bean
A	Castrol bean + maize, 1:1 ratio	1,080	2,170	—
B	Castrol bean + maize, 1:2 ratio	728	2,290	—
C	Castrol bean + maize, 1:3 ratio	847	1,936	—
D	Castrol bean + bean, 1:2 ratio	1,148	—	1,005
E	Castrol bean + bean, 1:3 ratio	888	—	1,150
F	Castrol bean + bean, 1:4 ratio	746	—	858
G	Castrol bean + maize + bean, 1:2:1 ratio	561	1,991	92
H	Castrol bean + maize + bean, 1:1:2 ratio	1,242	1,449	540
I	Castrol bean + maize + bean, 1:2:2 ratio	697	1,657	340
J	Castrol bean alone	1,871	—	—
K	Maize alone	—	2,167	—
L	Bean alone	—	—	1,007

The cultivar used for castrol beans was *Amarela*, for maize the cultivar was *Maia* 4, and for beans it was *Vagem roxa*. Treatments A, D, and H resulted in a reduction of yields of about one-third of the castrol bean sole crop. Treatments A, B, C, and G resulted in about the same maize yields as obtained for the sole crop maize. Bean treatments D, E, and F had yields equal to the sole crop beans, whereas treatment G reduced bean yield to less than 10% of the sole crop. The higher proportions of maize and beans in mixtures with castrol beans reduced castrol bean yields considerably.

Note that the proportions of the various crops change in the example. It is important to understand that the densities per hectare for one experimental unit do not change, but only the proportion of the crops in a mixture changes. Thus, this example is in the spirit of this chapter, viz. the density per hectare is constant.

Example 13.3. The following experiment involved intercropping four crops, sorghum (S), cowpeas (C), maize (M), and beans (B), but only in pairs and as sole crops. It was carried out in 1974 at two locations, Caruaru and Serra Talhadá, Pernambuco, Brazil. The treatment design and yields in kg/ha are

	Yield (kg/ha) Caruaru				Serra Talhadá			
	S	B	C	M	S	B	C	M
S=sorghum (sole)	2,210	—	—	—	2,919	—	—	—
C=cowpeas (sole)	—	—	1,220	—	—	—	944	—
B=beans (sole)	—	69	—	—	—	984	—	—
M=maize (sole)	—	—	—	2,451	—	—	—	2,828
S + C	2,577	—	621	—	2,766	—	524	—
S + B	2,194	33	—	—	2,602	488	—	—
M + C	—	—	419	2,388	—	—	249	2,530
M + B	—	30	—	2,599	—	257	—	2,759

The yield of sorghum was relatively unaffected by adding cowpeas and was only slightly decreased by adding beans. Hence, any additional yield obtained from beans and cowpeas was a bonus. The yields of cowpeas and beans were affected more by maize than by sorghum, possibly due to the later maturity of sorghum and greater demands on water and nutrients after beans and cowpeas had matured more. The maize yields, whether in mixture or sole, were relatively the same. This is what was observed for cultivar *X* in Example 2.1 of Volume I. Hence, any additional bean and cowpea yields represent a bonus from intercropping with maize.

Example 13.4. This experiment involved five crops, a cotton cultivar=C, a maize cultivar=M, a bean cultivar=B, a sorghum cultivar=S, and two cultivars of palm, Capim Buffel=PB, and Capim Colonião=PC in four treatment combinations both with and without fertilizer. Ten blocks of a randomized complete block design were used. The experimental unit size was 2.1 m by 1 m (area harvested). The four treatments were

a: cotton (C) as a sole crop for 4 years;
b: C + M + B for 4 years;
c: C + M + B in year 1, Capim Buffel planted in year 2 and harvested in years 3 and 4;
d: C + S + B in year 1, Capim Colonião planted in year 2 and harvested in years 3 and 4.

The yields obtained were

	Yield (kg/ha), 1969-72 period					
	C	M	B	S	PB	PC
Treatment	without fertilizer					
a	2,461	—	—	—	—	—
b	2,163	1,415	280	—	—	—
c	2,277	1,415	248	—	245	—
d	1,784	—	185	2,232	—	18t.*
	C	M	B	S	PB	PC
	with fertilizer					
a	2,564	—	—	—	—	—
b	2,630	1,390	290	—	—	—
c	2,488	1,387	268	—	267	—
d	2,459	—	142	2,458	—	20t.*

* tons per hectare.

If the value of cotton is 7.00, the value of maize is 1.50, the value of beans is 5.00, the value of sorghum is 1.30, the value of PC is 0.025 cruzerios per kilogram, and the value of PB is 25.00/ton, we may compute the values for the various treatments.

Treatment		No fertilizer value	(%) of a	Fertilizer value	(%) of a
a–	cotton	17,227	100	17,948	100
b–	cotton	15,141		18,410	
	maize	2,112		2,085	
	bean	1,400		1,450	
	Total	18,653	108	21,945	122
c–	cotton	15,939		17,416	
	maize	2,122		2,080	
	bean	1,240		1,340	
	PB	6		7	
	Total	19,307	112	20,843	116
d–	cotton	12,488		17,213	
	sorghum	2,902		3,195	
	bean	925		710	
	PC	450		500	
	Total	16,765	97	21,618	120

Thus, the crop value was −3% to 22% higher for mixtures than for cotton grown alone. Treatment d was somewhat lower than the others on the nonfertilized plots. For the fertilized plots, cotton yields in mixtures were approximately equal to the sole crop yields. On the nonfertilized plots, the cotton yields were somewhat lower in the mixtures. Treatment d gave lower bean yields than treatments b and c whether fertilized or not.

As can be seen from the above examples, treatment designs can be many and varied. It is a good rule to include sole crops for all crops unless the crop is infrequently or never grown as a sole crop. Even if this were the case, it still might be wise to include all sole crops for biological modeling purposes and to have a point of reference. In the above examples, the ratio of the sole crop maize yield to bean yield was 2:1 in Example 13.2 and was 350:1 and 3:1 in the two locations of Example 13.3. The 350:1 ratio appears to be an abnormality, as bean yield was unduly low. The 2:1 and 3:1 ratios are somewhat smaller than the ratio in Examples 2.1 and 3.1 of Volume I. This again demonstrates the need to obtain yields for crops outside the experiment; land equivalent ratios using experimental values are highly suspect, as demonstrated in Example 13.3 where ratios of 350:1 and 3:1 were obtained at the two locations.

The treatments to be included in an experiment require thoughtful consideration. The goals for each intercropping experiment should be carefully pondered. Then, the treatment design can be formulated to meet the needs of the experiment.

13.3 Response Model Equations and Analyses

For the following discussion, we shall consider the situation where the treatments are in a randomized complete block design and are included only once in each block. The results are easily generalizable to other experiment designs; the following equations may be considered to be appropriate for the $v = 3$ sole crop yields:

$$\text{Crop one} \qquad Y_{1hi} = \mu_1 + \rho_{1h} + \tau_{1i} + \epsilon_{1hi}, \tag{13.1}$$

$$\text{Crop two} \qquad Y_{2hi} = \mu_2 + \rho_{2h} + \tau_{2i} + \epsilon_{2hi}, \tag{13.2}$$

$$\text{Crop three} \qquad Y_{3hi} = \mu_3 + \rho_{3h} + \tau_{3i} + \epsilon_{3hi}, \tag{13.3}$$

where Y_{fhi} is the response for the fth crop, $f = 1, 2, 3 = v$, in the hth block, $h = 1, 2, \ldots, r$, for the ith line of crop f, $i = 1, 2, \ldots, c_f$, μ_f is an overall mean effect for crop f, ρ_{fh} is the hth block effect for crop f, τ_{fi} is the ith line effect for crop f, and ϵ_{fhi} is a normal independent random variable with mean zero and variance $\sigma^2_{f\epsilon}$. A straightforward extension results in v equations for the v sole crop responses.

The following response model equations may be appropriate for mixtures of lines of three main crops. Generalization to v main crops is straightforward. The crops are assumed to be in a 1:1:1 ratio, i.e., one-third of the area for a sole crop would be devoted to each crop. Certain crops might have an equal number of plants/ha as well as equal areas. The response model equations for mixtures of three crops are

$$\begin{aligned} Y_{1hi(jg)} = {} & (\mu_1 + \rho_{1h} + \tau_{1i} + \delta_{1i})/3 + 2(\gamma_{i(j)} + \gamma_{i(g)})/3 \\ & + \pi_{i(jg)} + \epsilon_{1hi(jg)}, \end{aligned} \tag{13.4}$$

$$\begin{aligned} Y_{2h(i)j(g)} = {} & (\mu_2 + \rho_{2h} + \tau_{2j} + \delta_{2j})/3 + 2(\gamma_{(i)j} + \gamma_{j(g)})/3 \\ & + \pi_{(i)j(g)} + \epsilon_{2h(i)j(g)}, \end{aligned} \tag{13.5}$$

and

$$\begin{aligned} Y_{3h(ij)g} = {} & (\mu_3 + \rho_{3h} + \tau_{3g} + \delta_{3g})/3 + 2(\gamma_{(i)g} + \gamma_{(j)g})/3 \\ & + \pi_{(ij)g} + \epsilon_{3h(ij)g}, \end{aligned} \tag{13.6}$$

where $Y_{1hi(jg)}$ is the response for the ith line of main crop one in the hth block and in the mixture ijg, $Y_{2h(i)j(g)}$ and $Y_{3h(ij)g}$ are similarly defined; μ_f, ρ_{fh}, and τ_{fi} are defined above; δ_{fi} is the general mixing effect for the ith line of the fth main crop; $\gamma_{i(j)}$ is a bi-specific mixing effect for the ith line of crop one with the jth line of crop two with $\gamma_{(i)j}$ being that part of the interaction for line j of crop two; $\gamma_{i(g)}$, $\gamma_{j(g)}$, and $\gamma_{(j)g}$ are similarly defined; $\pi_{i(jg)}$ is the tri-specific mixing effect peculiar to line i of crop one in the mixture ijg; $\pi_{(i)j(g)}$ and $\pi_{(ij)g}$ are similarly defined; and $\epsilon_{1hi(jg)}$, $\epsilon_{2h(i)j(g)}$, and $\epsilon_{3h(ij)g}$ are normal random variables with means of zero and variances $\sigma^2_{1\epsilon 3}$, $\sigma^2_{2\epsilon 3}$, and $\sigma^2_{3\epsilon 3}$, respectively.

If the three main crops are in the proportions $p_1 : p_2 : p_3$ on an area basis, where $p_1 + p_2 + p_3 = 1$, then equations (13.4) to (13.6) may be rewritten as follows when $p_1 \geq p_2 \geq p_3$:

$$\begin{aligned} Y_{1hi(jg)} = {} & p_1(\mu_1 + \rho_{1h} + \tau_{1i} + \delta_{1i}) + 2p_2\gamma_{i(j)} + 2p_3\gamma_{i(g)} \\ & + 3p_3\pi_{i(jg)} + \epsilon_{1hi(jg)}, \end{aligned} \tag{13.7}$$

$$Y_{2h(i)j(g)} = p_2(\mu_2 + \rho_{2h} + \tau_{2j} + \delta_{2j}) + 2p_2\gamma_{(i)j} + 2p_3\gamma_{j(g)} + 3p_3\pi_{(i)j(g)} + \epsilon_{2h(i)j(g)}, \tag{13.8}$$

and

$$Y_{3h(ij)g} = p_3(\mu_3 + \rho_{3h} + \tau_{3g} + \delta_{3g}) + 2p_3\gamma_{(i)g} + 2p_3\gamma_{(j)g} + 3p_3\pi_{(ij)g} + \epsilon_{3h(ij)g}. \tag{13.9}$$

Interaction effects like $\gamma_{i(j)}$, $\gamma_{(i)j}$, etc. are defined where each line i and j have equal amounts of material to interact. Since the proportion of line i of crop one and line j of crop two only have $p_1 + p_2$ of the space in the experimental unit (e.u.) and since $p_1 \geq p_2$, there is only $2p_2$ of the e.u. available for interaction of equal amounts of material. For the three-factor interaction which is defined for a 1:1:1 ratio, there is only $3p_3$ of the e.u. usable for the three-factor interaction effect for $p_1 \geq p_2 \geq p_3$. This method of defining an interaction is consistent and unchanging for any set of relative proportions p_1: p_2: p_3 and for a changing number of crops in the mixture. A desirable feature of this definition of interaction is to free it of confounding with the variable density per hectare. Equations (13.4) to (13.6) and (13.7) to (13.9) may be extended in a straightforward manner for more than three crops in a mixture.

For equations (13.1) to (13.3), c_f, the number of lines of the fth crop, must be at least two. Alternatively, if $c_f = 1$, then the line must appear at least twice in at least some of the blocks in order to obtain an estimate of $\sigma^2_{f\epsilon}$. Given that $c_f \geq 2$ and that each treatment occurs once in each block, the solutions for restraints $\Sigma_h \hat{\rho}_{fh} = \Sigma_i \hat{\tau}_{fi} = 0$ are

$$\hat{\mu}_f = \sum_{h=1}^{r} \sum_{i=1}^{c_f} Y_{fhi}/rc_f = \bar{y}_{f\cdot\cdot}, \tag{13.10}$$

$$\hat{\rho}_{fh} = \bar{y}_{fh\cdot} - \bar{y}_{f\cdot\cdot}, \tag{13.11}$$

$$\hat{\tau}_{fi} = \bar{y}_{f\cdot i} - \bar{y}_{f\cdot\cdot}, \tag{13.12}$$

and

$$\hat{\sigma}^2_{f\epsilon} = \left[\sum_{h=1}^{r} \sum_{i=1}^{c_f} Y^2_{fhi} - \sum_{h=1}^{r} \frac{Y^2_{fh\cdot}}{c_f} - \sum_{i=1}^{c_f} \frac{Y^2_{f\cdot i}}{r} + \frac{Y^2_{f\cdot\cdot}}{rc_f}\right] \Big/ (r-1)(c_f-1). \tag{13.13}$$

Solutions for the ith line of crop one, the jth line of crop two, and the gth line of crop three for the response model equations (13.4) to (13.6) and subject to the additional restrictions that $\Sigma_{j=1}^{c_2} \hat{\gamma}_{i(j)} = \Sigma_{i=1}^{c_1} \hat{\gamma}_{(i)j} = \Sigma_{g=1}^{c_3} \hat{\gamma}_{i(g)} = \Sigma_{g=1}^{c_3} \hat{\gamma}_{j(g)} = \Sigma_{i=1}^{c_1} \hat{\gamma}_{(i)g} = \Sigma_{j=1}^{c_2} \hat{\gamma}_{(j)g} = \Sigma_j \Sigma_g \hat{\pi}_{i(jg)} = \Sigma_i \Sigma_j \hat{\pi}_{(ij)g} = \Sigma_i \Sigma_j \hat{\pi}_{(ij)g} = 0$ for each f and for the ijg combination of lines of the three crops are given as follows for mixtures of three crops, $f = 1, 2, 3$:

$$\hat{\delta}_{1i} = 3\bar{y}_{1\cdot i(\cdot\cdot)} - \bar{y}_{1\cdot i}, \tag{13.14}$$

$$\hat{\delta}_{2j} = 3\bar{y}_{2\cdot(\cdot)j(\cdot)} - \bar{y}_{2\cdot j}, \tag{13.15}$$

$$\hat{\delta}_{3g} = 3\bar{y}_{3\cdot(\cdot\cdot)g} - \bar{y}_{3\cdot g}, \tag{13.16}$$

$$\hat{\gamma}_{i(j)} = 3(v-2)[\bar{y}_{1\cdot i(j\cdot)} - \bar{y}_{1\cdot i(\cdot\cdot)}]/2(v-3), \tag{13.17}$$

$$\hat{\gamma}_{i(g)} = 3(v-2)[\bar{y}_{1\cdot i(\cdot g)} - \bar{y}_{1\cdot i(\cdot\cdot)}]/2(v-3), \tag{13.18}$$

$$\hat{\gamma}_{(i)j} = 3(v-2)[\bar{y}_{2\cdot(i)j(\cdot)} - \bar{y}_{2\cdot(\cdot)j(\cdot)}]/2(v-3), \tag{13.19}$$

$$\hat{\gamma}_{j(g)} = 3(v-2)[\bar{y}_{2\cdot(\cdot)j(g)} - \bar{y}_{2\cdot(\cdot)j(\cdot)}]/2(v-3), \tag{13.20}$$

$$\hat{\gamma}_{(i)g} = 3(v-2)[\bar{y}_{3\cdot(i\cdot)g} - \bar{y}_{3\cdot(\cdot\cdot)g}]/2(v-3), \tag{13.21}$$

$$\hat{\gamma}_{(j)g} = 3(v-2)[\bar{y}_{3\cdot(\cdot j)g} - \bar{y}_{3\cdot(\cdot\cdot)g}]/2(v-3), \tag{13.22}$$

$$\hat{\pi}_{i(jg)} = \bar{y}_{1\cdot i(jg)} - \left(\frac{v-2}{v-3}\right)\left(\bar{y}_{1\cdot i(j\cdot)} + \bar{y}_{1\cdot i(\cdot g)}\right) + \left(\frac{v-1}{v-3}\right)\bar{y}_{1\cdot i(\cdot\cdot)}, \tag{13.23}$$

$$\begin{aligned}\hat{\pi}_{(i)j(g)} = \bar{y}_{2\cdot(i)j(g)} &- \left(\frac{v-2}{v-3}\right)\left(\bar{y}_{2\cdot(i)j(\cdot)} + \bar{y}_{2\cdot(\cdot)j(g)}\right)\\ &+ \left(\frac{v-1}{v-3}\right)\bar{y}_{2\cdot(\cdot)j(\cdot)},\end{aligned} \tag{13.24}$$

$$\hat{\pi}_{(ij)g} = \bar{y}_{3\cdot(ij)g} - \left(\frac{v-2}{v-3}\right)\left(\bar{y}_{3\cdot(i\cdot)g} + \bar{y}_{3\cdot(\cdot j)g}\right) + \left(\frac{v-1}{v-3}\right)\bar{y}_{3\cdot(\cdot\cdot)g}, \tag{13.25}$$

$$\begin{aligned}\hat{\sigma}^2_{1\epsilon 3} = &\left[\sum_h\sum_i\sum_j\sum_g Y^2_{1hi(jg)} - \sum_h \frac{Y^2_{1h(\cdot\cdot)}}{c_1c_2c_3}\right.\\ &\left.-\sum_i\sum_j\sum_g \frac{Y^2_{1\cdot i(jg)}}{r} + \frac{Y^2_{1\cdot\cdot(\cdot\cdot)}}{rc_1c_2c_3}\right]\Big/ (r-1)(c_1c_2c_3-1),\end{aligned} \tag{13.26}$$

$$\begin{aligned}\hat{\sigma}^2_{2\epsilon 3} = &\left[\sum_h\sum_i\sum_j\sum_g Y^2_{2h(i)j(g)} - \sum_h \frac{Y^2_{2h(\cdot)\cdot(\cdot)}}{c_1c_2c_3}\right.\\ &\left.-\sum_i\sum_j\sum_g \frac{Y^2_{2\cdot(i)j(g)}}{r} + \frac{Y^2_{2\cdot(\cdot)\cdot(\cdot)}}{rc_1c_2c_3}\right]\Big/ (r-1)(c_1c_2c_3-1),\end{aligned} \tag{13.27}$$

$$\begin{aligned}\hat{\sigma}^2_{3\epsilon 3} = &\left[\sum_h\sum_i\sum_j\sum_g Y^2_{3h(ij)g} - \sum_h \frac{Y^2_{3h(\cdot\cdot)\cdot}}{c_1c_2c_3}\right.\\ &\left.-\sum_i\sum_j\sum_g \frac{Y^2_{3\cdot(ij)g}}{r} + \frac{Y^2_{3\cdot(\cdot\cdot)\cdot}}{rc_1c_2c_3}\right]\Big/ (r-1)(c_1c_2c_3-1),\end{aligned} \tag{13.28}$$

where $c_1, c_2, c_3 > 1$ for the last three equations and where the 3 in $\hat{\sigma}^2_{f\epsilon 3}$ refers to mixtures of three. Note that for v crops in mixtures of three the factor $c_1c_2c_3$ is replaced by $c_1c_2c_3(v-1)(v-2)/2$ in equations (13.26)–(13.28), to account for the

TABLE 13.1. Analysis of Variance for Crop One Responses as a Sole Crop and in a Mixture of Three Cultivars for Response Model Equations (13.1) and (13.4) for $v = 3$.

Source	Degrees of Freedom	Sum of squares
Total	$rc_1c_2c_3$	$\sum_{h=1}^{r}\sum_{i=1}^{c_1}\sum_{j=1}^{c_2}\sum_{g=1}^{c_3} Y_{1hi(jg)}^2 + \sum_h\sum_i Y_{1hi}^2$
Correction for mean	1	$(Y_{1\cdot\cdot(\cdot\cdot)} + Y_{1\cdot\cdot})^2/(rc_1c_2c_3 + rc_1) = C$
Blocks	$r-1$	$\sum_h Y_{1h\cdot(\cdot\cdot)}^2/(c_1c_2c_3 + c_1) - C$
Treatments	$c_1c_2c_3 + c_1 - 1$	$\sum_i\sum_j\sum_g Y_{1\cdot i(jg)}^2/r + \sum_i Y_{1\cdot\cdot}^2/rc_1 - C$
Sole crop one lines	$c_1 - 1$	$\sum_i Y_{1\cdot i}^2/r - Y_{1\cdot\cdot}^2/rc_1$
Sole crop vs. mix. Of 3	1	$(Y_{1\cdot\cdot}^2/rc_1) + (Y_{1\cdot\cdot(\cdot\cdot)}^2/rc_1c_2c_3) - C$
General mixing effects	$c_1 - 1$	$rc_2c_3\sum_i \hat{\delta}_{1i}/3\hat{\delta}_{1i}^2$
Two-factor interaction of crops one and two	$c_1(c_2-1)$	$2rc_3\sum_i\sum_j \hat{\gamma}_{i(j)}^2/3$
Two-factor interaction of crops one and three	$c_1(c_3-1)$	$2rc_2\sum_i\sum_g \hat{\gamma}_{i(g)}^2/3$
Three-factor interaction	$c_1(c_2-1)(c_3-1)$	$r\sum_i\sum_j\sum_g \hat{\pi}_{i(jg)}^2$
Blocks × treatments	$(r-1)(c_1c_2c_3 + c_1 - 1)$	By subtraction
Blocks × sole crops	$(r-1)(c_1 - 1)$	Formula (13.13)
Blocks × mixtures	$(r-1)(c_1c_2c_3 - 1)$	Formula (13.26)
Blocks × soles vs. mixtures	$r-1$	By subtraction

total number of mixtures involving crop f and its c_f cultivars. Similar solutions are obtainable for equations (13.7), (13.8), and (13.9). The solutions involve the proportions p_1, p_2, and p_3 with the above means. Note that $\hat{\rho}_{fh}$ from both sole crops and mixtures for crop f is obtained as

$$\hat{\rho}_{fh} = \left(\sum_{i=1}^{v} Y_{fhi} + \sum_i\sum_j\sum_g Y_{fhi(jg)}\right)\Bigg/ (c_f + c_1c_2c_3) - \left(\sum_h\sum_i Y_{fhi} + \sum_h\sum_i\sum_j\sum_g Y_{fhi(jg)}\right)\Bigg/ r(c_f + c_1c_2c_3). \quad (13.29)$$

An analysis of variance table for the responses for one crop is given in Table 13.1. The computations for the various sums of squares are straightforward and involve only totals and solutions for effects. Similar analysis of variance tables may be set up for other crops. The variance of a difference between two sole crop one-line effects, say $\hat{\tau}_{1i} - \hat{\tau}_{1i'} = \bar{y}_{1\cdot i} - \bar{y}_{1\cdot i'}$, is obtained from formula (13.13) as $2\hat{\sigma}^2_{1\epsilon}/r = \hat{V}(\hat{\tau}_i - \hat{\tau}_{i'})$, $i \neq i'$. The estimated variance of a difference between two mixture means, $\hat{y}_{1\cdot i(jg)} - \bar{y}_{1\cdot i'(jg)}$, is $\hat{V}(\bar{y}_{1\cdot i(jg)} - \bar{y}_{1\cdot i'(jg)}) = 2\hat{\sigma}^2_{1\epsilon 3}/r$, and for the two means $\bar{y}_{1\cdot i(j\cdot)} - \bar{y}_{1\cdot i'(j\cdot)}$ it is $2\hat{\sigma}^2_{1\epsilon 3}/rc_1$. It should be noted that the variances of $Y_{fhi(jg)}$ may not all be estimates of the same single parameter $\sigma^2_{1\epsilon 3}$. The discussion here is based on the assumption of homoscedastic variances. The sole crop one mean may be compared with the mean of three crop mixture responses for crop one as follows:

$$t = \left(\bar{y}_{1\cdot\cdot} - \bar{y}_{1\cdot\cdot(\cdot\cdot)}\right) \Big/ \sqrt{\frac{\hat{\sigma}^2_{1\epsilon}}{rc_1} + \frac{\hat{\sigma}^2_{1\epsilon 3}}{rc_1c_2c_3}}$$

and compared with a computed t value equal to

$$t'_\alpha = \left(t_{\alpha,f_1}\frac{\hat{\sigma}^2_{1\epsilon}}{rc_1} + t_{\alpha,f_{13}}\frac{\hat{\sigma}^2_{1\epsilon 3}}{rc_1c_2c_3}\right) \Big/ \left(\frac{\hat{\sigma}^2_{1\epsilon}}{rc_1} + \frac{\hat{\sigma}^2_{1\epsilon 3}}{rc_1c_2c_3}\right),$$

where $f_1 = (r-1)(c_1-1)$ and $f_{13} = (r-1)(c_1c_2c_3-1)$ are the degrees of freedom associated with $\hat{\sigma}^2_{\epsilon 1}$ and $\hat{\sigma}^2_{1\epsilon 3}$, respectively. This is the Cochran formula for approximating a tabulated t-statistic at the α percent level for the Behrens–Fisher situation. (See, e.g., Grimes and Federer, 1984.) Also, in the above it was assumed that there were c_1 lines of crop one, c_2 lines of crop two, and c_3 lines of crop three in all possible combinations. This means that there were $rc_1c_2c_3$ yields entering into the mean $\bar{y}_{1\cdot\cdot(\cdot\cdot)}$.

Example 13.5. For the experiment described in Section 12.4, the individual plant dry weights of the six secondary crops,

A	wild oat	D	birdsfoot trefoil
B	coriander	E	alfalfa
C	lentils	F	chamomile

were available and are given in Table 13.2. We shall use these data to demonstrate an analysis for six main crops as sole crops, in mixtures of three, and in a mixture of six, all overseeded with barley. Here, $c_{fi} = 1$ for f =A, . . . ,F, and there are $(v-1)(v-2)/2 = 10$ combinations where crop i occurs. There were 12 plants entering the total yield for sole crops, 4 plants of each crop for mixtures of 3, and 2 plants for mixtures of all 6 cultivars. The words crop and cultivar are used interchangeably in this chapter. The number of plants in an experimental unit for these responses was a constant number, 12 of 1 or more of these 6 cultivars and 20 barley plants.

For the combination ABE in block one, there were two missing plants for A. To bring the total biomass to a four-plant total, the yield of the two plants was

TABLE 13.2. Biomass in Grams for Single Cultivar and Mixtures of Three Cultivars, ijg.

Treatment	Block 1			Block 2			Block 3			Totals		
ijg	i	j	g	i	j	g	i	j	g	i	j	g
A	24.04			15.87			27.84			67.75		
B	2.21			1.58			1.12			4.91		
C	47.98			41.76			28.89			118.63		
D	4.02			1.09			3.04			8.15		
E	13.72			4.81			8.37			26.90		
F	3.34			0.49			0.95			4.78		
ABC	6.54	0.95	28.53	7.05	0.74	14.55	0.29	0.39	9.46	13.88	2.08	52.54
ABD	6.13	0.61	1.63	9.78	0.49	0.31	11.84	0.71	0.47	27.75	1.81	2.41
ABE	3.96	1.70	4.17	10.40	0.41	2.26	6.14	0.93	3.08	20.50	3.04	9.51
ABF	5.12	0.63	0.45	5.88	0.40	0.03	7.47	0.64	0.28	28.47	1.67	0.76
ACD	6.79	13.97	1.51	3.50	13.72	0.34	5.52	16.58	0.83	15.81	44.27	2.68
ACE	8.94	9.03	0.93	5.65	25.07	2.21	7.23	12.06	1.42	21.82	46.16	4.56
ACF	9.03	23.07	1.37	5.78	11.89	0.18	7.65	18.06	1.03	22.46	53.02	2.58
ADE	7.19	0.72	3.37	6.08	0.40	1.84	7.89	0.55	1.99	21.16	1.67	7.20
ADF	9.45	1.08	1.11	2.16	0.77	0.32	9.07	0.22	0.23	20.68	2.07	1.66
AEF	3.93	7.58	1.26	7.77	4.58	0.06	12.33	4.28	4.47	24.03	16.44	1.79
BCD	1.35	12.27	0.89	0.55	10.89	0.25	0.68	10.16	0.42	2.58	33.32	1.56
BCE	1.72	22.28	2.78	1.09	23.26	4.57	1.19	15.21	1.84	4.00	60.75	9.19
BCF	1.26	20.64	1.15	0.67	15.23	0.44	1.04	18.16	0.49	2.97	54.03	2.08
BDE	0.85	0.69	2.65	1.20	1.08	7.21	0.68	0.28	1.15	2.73	2.05	11.01
BDF	0.68	0.18	0.30	0.46	0.03	0.06	0.48	0.26	0.08	1.62	0.47	0.44
BEF	0.52	1.57	0.28	0.51	5.22	0.32	1.20	1.82	0.48	2.23	8.61	1.08
CDE	19.53	0.54	2.62	14.52	0.28	1.12	15.79	1.03	6.89	49.84	1.85	10.63
CDF	13.45	1.23	1.21	35.10	0.67	0.21	19.26	0.43	0.33	67.81	2.33	1.75
CEF	19.15	3.01	1.24	13.67	1.50	0.00	20.93	2.35	1.32	53.75	6.86	2.56
DEF	1.43	3.89	1.06	0.24	1.77	0.16	0.20	2.67	0.26	1.87	3.33	1.48
ABC	2.64	0.40	13.61	0.80	0.10	10.73	4.50	0.73	14.08	7.94	1.23	38.42
DEF	0.35	0.35	0.32	0.71	1.76	0.07	0.60	1.82	0.54	1.66	4.51	0.93

multiplied by 2 to obtain 3.96; that is, the mean of the remaining plants is used as the missing plant value. In block three, combination ACD had one missing value for A. Therefore, the total of three plants, 4.14, was multiplied by 4/3 to obtain 5.52. In block three, ABCDEF had one missing plant for C, and the one plant value was doubled to obtain 2(7.04) = 14.08. With such adjustments, all mixture yields are on either a two-plant or four-plant basis, depending on the combination. After performing the analyses, it was not noticed that there was also one missing plant for D in ADF of block three. The appropriate value should be (0.22)(4/3) = 0.27 rather than 0.22 as given in Table 13.3 for crop D. Since this value differs little from the appropriate value, the analyses were not redone but are left as an exercise for the reader. The procedure used would be a least squares solution for minimizing the among plants within cultivar and experimental unit mean square. Alternatively, an unequal numbers analysis could have been used. This would give a least squares solution minimizing the treatment by block sum of squares. Both procedures should give approximately the same results in this case.

Let us consider the following response model equations for the data from cultivar $i, i =$A,B,C,D,E,F, in this experiment. These are different but in the spirit of those given in (13.1) to (13.6). For sole crop $i =$A and effects on a *four-plant basis* [see Appendix 13A, equations (13.58) to (13.70)], the response equations are

$$Y_{hA} = \mu + \tau_A + \rho_{hA} + \epsilon_{hA}, \qquad h = 1, 2, \ldots, r, \tag{13.30}$$

$$\begin{aligned} Y_{hA(jg)} &= \mu + \tau_A + \rho_{hA} + \frac{1}{2}\left(\delta_{A(j)} + \delta_{A(g)}\right) + \pi_{A(jg)} \\ &\quad + \epsilon_{hA(jg)}, \qquad j, g = \text{B,C,D,E,F} \end{aligned} \tag{13.31}$$

and

$$Y_{hA(BCDEF)} = \mu + \tau_A + \rho_{hA} + 2\bar{\delta}_{A(\cdot)} + 2\beta_{A(BCDEF)} + \epsilon_{hA(BCDEF)}. \tag{13.32}$$

The above form for the parameters was used since all mixtures of three yields were on a four-plant basis. The sole crop yields for 12 plants were divided by 3 to put them on a four-plant basis and the yields from mixtures of 6 were multiplied by 2. These are the responses given in Table 13.3. The error variances should now be comparable and approximately equal.

The resulting normal equations under variance equality and model constraints $\Sigma_{h=1}^{r}\rho_{hi} = 0$, $\Sigma_{g=1,\neq j}^{v}\pi_{i(jg)} = 0$, and $\Sigma_{j=1}^{v}\delta_{i(j)} = (v-1)\hat{\delta}_{i(\cdot)}$ for $i \neq j$, $g = 1, 2, \ldots, v$ (on a four-plant basis) are given by equations (13.61) to (13.65). The solutions [see (13.66)–(13.70)] are

$$\hat{\mu} + \hat{\tau}_i = \bar{y}_{\cdot i}, \tag{13.33}$$

$$\hat{\delta}_{i(\cdot)} = \bar{y}_{\cdot i(\cdot\cdot)} - \bar{y}_{\cdot i}, \tag{13.34}$$

$$\hat{\delta}_{i(j)} = \frac{2}{3r} Y_{\cdot i(j\cdot)} - \frac{5}{3}\bar{y}_{\cdot i(\cdot\cdot)} - \bar{y}_{\cdot i}, \tag{13.35}$$

$$\hat{\pi}_{i(jg)} = \bar{y}_{\cdot i(jg)} - \frac{1}{3r}\left(Y_{\cdot i(j\cdot)} + Y_{\cdot i(g\cdot)}\right) + \frac{5}{3}\bar{y}_{\cdot i(\cdot\cdot)}, \tag{13.36}$$

$$2\hat{\beta}_{i(\text{all but } i)} = \bar{y}_{i(\text{all but } i)} - 2\bar{y}_{\cdot i(\cdot\cdot)} + \bar{y}_{\cdot i}. \tag{13.37}$$

For a random set of blocks, the variances for crop i [equations (13.76)–(13.80)] are

$$V(\hat{\mu} + \hat{\tau}_i) = \left(\sigma^2_{\epsilon i} + \sigma^2_{\rho i}\right)/r, \tag{13.38}$$

$$V(\hat{\delta}_{i(\cdot)}) = 11\sigma^2_{\epsilon i}/10r, \tag{13.39}$$

$$V(\hat{\delta}_{i(j)}) = 13\sigma^2_{\epsilon i}/6r, \tag{13.40}$$

$$V(\hat{\pi}_{i(jg)}) = \sigma^2_{\epsilon i}/2r, \tag{13.41}$$

and

$$V(\hat{\beta}_{i(\text{all but } i)}) = 12\sigma^2_{\epsilon i}/5r, \tag{13.42}$$

where $\sigma^2_{\rho i}$ is the variance component among ρ_{hi} and $\sigma^2_{\epsilon i}$ is a common variance for all error components in the design for cultivar i.

The data in Table 13.2 have been rearranged, put on a four-plot basis, and are given in Table 13.3 by cultivar. The application of the above formulas to the data of Table 13.3 is illustrated below for cultivar or crop A (results correct to rounding errors).

$$\hat{\mu} + \hat{\tau}_A = \frac{22.58}{3} = 7.527 = \bar{y}_{\cdot i},$$

$$\hat{\delta}_{A(\cdot)} = \frac{206.56}{3(10)} - \frac{22.58}{3} = 6.885 - 7.527 = -0.641,$$

$$\hat{\delta}_{A(B)} = \frac{2}{3(3)}(13.88 + 27.75 + 20.50 + 18.47)$$
$$- \frac{5}{3}(6.885) - 7.527 = -1.091,$$

$$\hat{\delta}_{A(C)} = \frac{2}{9}(73.97) - 19.002 = -2.564,$$

$$\hat{\delta}_{A(D)} = \frac{2}{9}(85.40) - 19.002 = -0.024,$$

$$\hat{\delta}_{A(E)} = \frac{2}{9}(87.51) - 19.002 = 0.445,$$

$$\hat{\delta}_{A(F)} = \frac{2}{9}(85.64) - 19.002 = 0.029,$$

$$\hat{\pi}_{A(BC)} = \frac{13.88}{3} - \frac{(80.60 + 73.97)}{3(3)} + \frac{5}{3}(6.8853) = -1.072,$$

$$\hat{\pi}_{A(BD)} = \frac{27.75}{3} - \frac{166.00}{9} + 11.476 = 2.281,$$

$$\hat{\pi}_{A(BE)} = \frac{20.50}{3} - \frac{168.11}{9} + 11.476 = -0.370,$$

$$\hat{\pi}_{A(BF)} = \frac{18.47}{3} - \frac{166.24}{9} + 11.476 = -0.838,$$

$$\hat{\pi}_{A(CD)} = \frac{15.81}{3} - \frac{159.37}{9} + 11.476 = -0.962,$$

$$\hat{\pi}_{A(CE)} = \frac{21.82}{3} - \frac{161.48}{9} + 11.476 = 0.807,$$

$$\hat{\pi}_{A(CF)} = \frac{22.46}{3} - \frac{159.61}{9} + 11.476 = 1.228,$$

$$\hat{\pi}_{A(DE)} = \frac{21.16}{3} - \frac{172.91}{9} + 11.476 = -0.683,$$

$$\hat{\pi}_{A(DF)} = \frac{20.68}{3} - \frac{171.04}{9} + 11.476 = -0.635,$$

$$\hat{\pi}_{A(EF)} = \frac{24.03}{3} - \frac{173.15}{9} + 11.476 = 0.247,$$

$$\hat{\beta}_{A(BCDEF)} = \frac{1}{2}\left(\frac{15.88}{3} - \frac{206.56}{15} + \frac{22.58}{3}\right) = -0.475.$$

The remaining values are computed in a similar manner and are presented in Table 13.4.

Analyses of variance for each of the six cultivar yields are presented in Table 13.5. Standard computational procedures are used for all but the last two lines of an analysis of variance table. These are the variations among general mixing effects, the $\hat{\delta}_{i(j)}$ values, and the variation among the $\hat{\pi}_{i(jg)}$ interaction or specific mixing effect values. A few of the computations are left as an exercise for the reader. For cultivar A, the sum of squares is computed as

$$\frac{3r}{4}\left\{\sum_j \hat{\delta}^2_{A(j)} - \left(\sum \hat{\delta}_{A(j)}\right)^2 \Big/ (v-1)\right\}$$

$$= \frac{3(3)}{4}\left\{(-1.091)^2 + (-2.564)^2 + (-0.024)^2 + 0.445^2 + 0.029^2 \right.$$

$$\left. - \frac{1}{5}(-1.091 + \cdots + 0.029)^2\right\}$$

$$= \frac{9}{4}\{7.964 - 2.054\}$$

$$= 13.298 \text{ with } v - 2 = 4 \text{ degrees of freedom.}$$

TABLE 13.3. Plant Dry Weights by Cultivar

Cultivar A yields (g; 4-plant dry weight)					Cultivar B yields (g; 4-plant dry weight)				
Treatment	Block				Treatment	Block			
ijg	1	2	3	Total	*ijg*	1	2	3	Total
A/3	8.01	5.29	9.28	22.58	B/3	0.74	0.53	0.37	1.64
ABC	6.54	7.05	0.29	13.88	ABC	0.95	0.74	0.39	2.08
ABD	6.13	9.78	11.84	27.75	ABD	0.61	0.49	0.71	1.81
ABE	3.96	10.40	6.14	20.50	ABE	1.70	0.41	0.93	3.04
ABF	5.12	5.88	7.47	18.47	ABF	0.63	0.40	0.64	1.67
ACD	6.79	3.50	5.52	15.81	BCD	1.35	0.55	0.68	2.58
ACE	8.94	5.65	7.23	21.82	BCE	1.72	1.09	1.19	4.00
ACF	9.03	5.78	7.65	22.46	BCF	1.26	0.67	1.04	2.97
ADE	7.19	6.08	7.89	21.16	BDE	0.85	1.20	0.68	2.73
ADF	9.45	2.16	9.07	20.68	BDF	0.68	0.46	0.48	1.62
AEF	3.93	7.77	12.33	24.03	BEF	0.52	0.51	1.20	2.23
2(ABCDEF)	5.28	1.60	9.00	15.88	2(ABCDEF)	0.80	0.20	1.46	2.46
Total 3	67.08	64.05	75.43	206.56	Total 3	10.27	6.52	7.94	24.73
Total	80.37	70.94	93.71	245.02	Total	11.81	7.25	9.77	28.83

Cultivar C yields (g; 4-plant dry weight)					Cultivar D yields (g; 4-plant dry weight)				
Treatment	Block				Treatment	Block			
ijg	1	2	3	Total	*ijg*	1	2	3	Total
C/3	15.99	13.92	9.63	39.54	D/3	1.34	0.36	1.01	2.71
ABC	28.53	14.55	9.46	52.54	ABD	1.63	0.31	0.47	2.41
ACD	13.97	13.72	16.58	44.27	ACD	1.51	0.34	0.83	2.68
ACE	9.03	25.07	12.06	46.16	ADE	0.72	0.40	0.55	1.67
ACF	23.07	11.89	18.06	53.02	ADF	1.08	0.77	0.22	2.07
BCD	12.27	10.89	10.16	33.32	BCD	0.89	0.25	0.42	1.56
BCE	22.28	23.26	15.21	60.75	BDE	0.69	1.08	0.28	2.05
BCF	20.64	15.23	18.16	54.03	BDF	0.18	0.03	0.26	0.47
CDE	19.53	14.52	15.79	49.84	CDE	0.54	0.28	1.03	1.85
CDF	13.45	35.10	19.26	67.81	CDF	1.23	0.67	0.43	2.33
CEF	19.15	13.67	20.93	53.75	DEF	1.43	0.24	0.20	1.87
2(ABCDEF)	27.22	21.46	28.16	76.84	2(ABCDEF)	0.70	1.42	1.20	3.32
Total 3	181.92	177.90	155.67	515.49	Total 3	9.90	4.37	4.69	18.96
Total	225.13	213.28	193.46	631.87	Total	11.94	6.15	6.90	24.99

(continued)

TABLE 13.3. (continued)

Cultivar E yields (g; 4-plant dry weight)					Cultivar F yields (g; 4-plant dry weight)				
Treatment	Block				Treatment	Block			
ijg	1	2	3	Total	ijg	1	2	3	Total
E/3	4.57	1.60	2.79	8.96	F/3	1.11	0.16	0.32	1.59
ABE	4.17	2.26	3.08	9.51	ABF	0.45	0.03	0.28	0.76
ACE	0.93	2.21	1.42	4.56	ACF	1.37	0.18	1.03	2.58
ADE	3.37	1.84	1.99	7.20	ADF	1.11	0.32	0.23	1.66
AEF	7.58	4.58	4.28	16.44	AEF	1.26	0.06	0.47	1.79
BCE	2.78	4.57	1.84	9.19	BCF	1.15	0.44	0.49	2.08
BDE	2.65	7.21	1.15	11.01	BDF	0.30	0.06	0.08	0.44
BEF	1.57	5.22	1.82	8.61	BEF	0.28	0.32	0.48	1.08
CDE	2.62	1.12	6.89	10.63	CDF	1.21	0.21	0.33	1.75
CEF	3.01	1.50	2.35	6.86	CEF	1.24	0.00	1.32	2.56
DEF	3.89	1.77	2.67	8.33	DEF	1.06	0.16	0.26	1.48
2(ABCDEF)	1.86	3.52	3.64	9.02	2(ABCDEF)	0.64	0.14	1.08	1.86
Total 3	32.57	32.28	27.49	92.34	Total 3	9.43	1.78	4.97	16.18
Total	39.00	37.40	33.92	110.32	Total	11.18	2.08	6.37	19.63

Note: Sole crop weights divided by 3 and mixtures yields from mixture ABCDEF were multiplied by 2 to bring all yields to a 4-plant basis.

Note that $9\left(\sum \hat{\delta}_{i(j)}\right)^2/4(5)$ is not the sum of squares for mixtures of three crops versus a sole crop. However, this sum of squares multiplied by the factor $4v/r(n_1+n_2) = 24/99 = 8/33$ yields the sum of squares for mixtures of three versus a sole crop. n_1 is the number of experimental units for the sole crop, i.e., three, n_2 is the number of experimental units occupied by mixtures of three cultivars, i.e., 30, $r = 3$, and $v = 6$. Thus,

$$\left(\frac{9}{4}\right)\left(\frac{8}{33}\right)\frac{(\sum \hat{\delta}_{i(j)})^2}{5}$$

$$= \left(\frac{9}{4}\right)\left(\frac{8}{33}\right)(2.054) = \left(\frac{6}{11}\right)(2.054) = 1.120.$$

The sum of squares for the interaction of specific mixing effects, $\pi_{i(jg)}$, is computed as follows for intercrop A:

$$r\sum_{j<g}\sum \hat{\pi}^2_{A(jg)} = 3\{(-1.072)^2 + 2.281^2 + \cdots + 0.247^2\} = 33.620$$

with $(v-1)(v-4)/2 = 5$ degrees of freedom. These two sums of squares should add to the sum of squares among mixtures of three; i.e., $13.298 + 33.620 = 46.918$, which is equal to 46.924 within rounding errors.

If significance testing is done, it may be noted that there are only two mean squares significant at the 5% level. Those are the mean squares for among the estimated mixing effects $\hat{\delta}_{i(j)}$ for i = B and F. The corresponding F-statistics are

TABLE 13.4. Estimated Effects from Cultivar Yields in Table 13.3. Four-Plant Basis.

	Cultivar i (g)					
Effect	A	B	C	D	E	F
$\hat{\mu} + \hat{\tau}_i$	7.527	0.547	13.180	0.903	2.987	0.530
$\hat{\delta}_{i(\cdot)}$	−0.641	0.278	4.003	−0.271	0.091	0.009
$\hat{\delta}_{i(A)}$		−0.010	1.735	0.006	0.263	0.080
$\hat{\delta}_{i(B)}$	−1.091		2.768	−0.514	0.399	−0.460
$\hat{\delta}_{i(C)}$	−2.564	0.664		−0.086	−1.174	0.564
$\hat{\delta}_{i(D)}$	−0.024	0.021	1.568		0.143	−0.244
$\hat{\delta}_{i(E)}$	0.445	0.746	4.959	−0.303		0.107
$\hat{\delta}_{i(F)}$	0.029	−0.034	8.984	−0.459	0.826	
$\hat{\pi}_{i(AB)}$			2.082	0.154	−0.148	−0.087
$\hat{\pi}_{i(AC)}$		−0.181		0.030	−1.011	0.008
$\hat{\pi}_{i(AD)}$		0.051	−0.075		−0.790	0.106
$\hat{\pi}_{i(AE)}$		0.098	−1.141	−0.198		−0.027
$\hat{\pi}_{i(AF)}$		0.032	−0.866	0.013	1.949	
$\hat{\pi}_{i(BC)}$	−1.072			−0.083	0.464	0.111
$\hat{\pi}_{i(BD)}$	2.281		−4.242		0.412	−0.031
$\hat{\pi}_{i(BE)}$	−0.370		3.206	0.189		0.007
$\hat{\pi}_{i(BF)}$	−0.839		−1.046	−0.260	−0.729	
$\hat{\pi}_{i(CD)}$	−0.962	−0.029			1.072	−0.107
$\hat{\pi}_{i(CE)}$	0.807	0.082		−0.092		−0.012
$\hat{\pi}_{i(CF)}$	1.228	0.128		0.146	−0.526	
$\hat{\pi}_{i(DE)}$	−0.683	−0.021	0.169			0.032
$\hat{\pi}_{i(DF)}$	−0.636	−0.001	4.147		−0.694	
$\hat{\pi}_{i(EF)}$	−0.247	−0.159	−2.235	0.101		
$\hat{\beta}_{i(\text{all but } i)}$	−0.475	−0.141	3.880	0.373	−0.081	0.036

$0.345/0.105 = 3.29$ and $0.342/0.096 = 3.56$. The tabulated F-values at the 5% and 2.5% levels are $F_{.05}(4, 22) = 2.82$ and $F_{.025}(4, 22) = 3.44$.

It should be noted that there appear to be outliers in these data. A study of the residuals should be made before performing significance tests or computing confidence intervals. An example of a possible outlier is the value of 0.29 in Table 13.3 for cultivar A, block three, for the mixture ABC. For cultivar C, the values 25.07 for ACE in block two and 35.10 for CDF in block two appear to be the cause of large residuals resulting in a rather large blocks by treatment mean square and a large coefficient of variation (see Table 13.6). A study of the residuals as done in Chapter 12 is left as an exercise for the reader. Likewise, the data were not studied to determine whether or not a transformation was desirable. However, an analysis of variance was performed on the logarithms of the responses in Table 13.3. The F-ratios were computed and are given in Table 13.6 for the block and treatment mean squares. The F-ratios are roughly the same for yields and for logarithms of yields. The coefficients of variation must be considered as quite large. This reinforces the comment that a study of residuals should be done for

TABLE 13.5. Analyses of Variance for Data from Table 13.3, Four-Plant Basis.

Source of Variation	Degrees of freedom	Sum of squares A	Sum of squares B	Sum of squares C	Mean squares A	Mean squares B	Mean squares C
Total	36	1915.211	28.153	12,411.31			
CFM	1	1667.633	23.088	11,090.55			
Blocks = R	2	21.815	0.870	42.67	10.908	0.435	21.34
Treatment = T	11	55.536	1.879	516.54	5.049	0.171	46.96
Mixture of 3	9	46.924	1.668	260.16	5214	0.185	28.91
Sole vs. 3	1	1.122	0.210	43.70			
Rest vs. 6	1	7.490	0.001	212.68			
R × T	22	170.226	2.316	761.55	7.738	0.105	34.62
Mixture 3 × R	18	149.404	1.605	716.88	8.3000	0.089	39.83
Remainder	4	20.822	0.711	44.67	5.206	0.178	11.17
General	4	13.298	1.381	86.23	3.324	0.345	31.56
Interaction	5	33.620	0.287	173.94	6.727	0.057	34.79

Source of Variation	Degrees of freedom	Sum of squares D	Sum of squares E	Sum of squares F	Mean squares D	Mean squares E	Mean squares F
Total	36	24.312	438.702	17.789			
CFM	1	17.347	338.070	10.704			
Blocks = R	2	1.652	1.124	3.454	0.826	0.562	1.727
Treatment = T	11	1.865	30.011	1.521	0.170	2.728	0.138
Mixture of 3	9	1.107	29.978	1.502	0.123	3.499	0.167
Sole vs. 3	1	0.201	0.023	0.000			
Rest vs. 6	1	0.557	0.011	0.018			
R × T	22	3.447	69.496	2.110	0.157	3.159	0.096
Mixture 3 × R	18	2.403	62.978	1.651	0.133	3.331	0.092
Remainder	4	1.044	6.518	0.459	0.261	1.630	0.115
General	4	0.464	5.102	1.365	0.116	1.276	0.342
Interaction	5	0.643	24.282	0.137	0.129	4.856	0.027

these data. The investigator should be consulted to ascertain why there was so much variation. Better experiment design, appropriate statistical analyses, better experimental technique, and/or larger experimental units may be needed. If these do not solve the problem of excessive variation, then the only recourse is to increase the number of replicates.

Various variances may be computed using equations (13.38) to (13.42). Then, comparisonwise confidence intervals may be computed. These are given in Table 13.7. The estimated variances for the estimated effects for cultivar A are

$$V(\mu + \hat{\tau}_A) = [7.738 + (10.908 - 7.738)/12]/r = 2.6674,$$

$$V(\hat{\delta}_{A(\cdot)}) = 11(7.738)/10r = 2.8373,$$

$$V(\hat{\delta}_{A(j)}) = 13(7.738)/6r = 5.5886,$$

TABLE 13.6. F-Ratios and Coefficients of Variation for Data of Table 13.3, Using Original Responses and Logarithm of Responses. Log(yield + 1) was used in place of log(yield).

		Cultivar (F-ratio)					
		A	B	C	D	E	F
Yields	Block	1.41	4.13	0.62	5.27	0.18	18.01
	Treatment	0.65	1.62	1.36	1.08	0.86	1.44
Log (yield)	Block	0.29	.	0.68	5.50	0.16	19.98
	Treatment	0.71	.	1.47	1.29	0.84	1.45

	Coefficient of variation (%)					
	A	B	C	D	E	F
Yield	41	41	34	57	58	57
Log (yield)	40	·	11	45	57	47

$$V(\hat{\pi}_{A(jg)}) = 7.738/2r = 1.2897,$$

and

$$V(\hat{\beta}_{A(BCDEF)}) = 12(7.738)/5r = 6.1904.$$

The variances for the remaining cultivars are computed in a similar manner and are presented in Table 13.7.

The 95% comparisonwise half-confidence intervals are computed by multiplying the square root of the estimated variance by the tabulated value for the t-statistic for $\alpha = .05$, and degrees of freedom equal to 22, i.e., $t_{.05,22} = 2.074$. Thus,

$$\begin{aligned}\hat{\delta}_{A(\cdot)} \pm t_{.05,22}\sqrt{V\left(\hat{\delta}_{A(\cdot)}\right)} &= -0.641 \pm 2.074\sqrt{2.8373} \\ &= -0.64 \pm 3.49 \text{ to two decimals}\end{aligned}$$

is the 95% confidence interval for $\delta_{A(\cdot)}$.

Until the problem of outliers is resolved, the following statements are to be considered tentative in interpreting the data and using significance tests. All of the $\delta_{i(\cdot)}$ confidence intervals contain the value zero and hence are not significantly different from zero. Two of the confidence intervals for $\delta_{B(j)}$, i.e., $\delta_{B(C)}$ and $\delta_{B(E)}$, do not contain the value zero; they are significantly different from zero at the 5% level. For cultivar F, one value, $\hat{\delta}_{F(C)}$, is significantly different from zero. For the $\hat{\pi}_{i(jg)}$, $\hat{\pi}_{A(BD)}$ is nearly significant and $\hat{\pi}_{E(AF)}$ is significant. Since there are 50 $\hat{\pi}_{i(jg)}$ values and if they were independent, which they are not, one would expect $.05(50) = 2.5$ to be significant based on random sampling. Since two of the $\hat{\pi}_{i(jg)}$ were significant, this is what would be expected based on random sampling. Again, the reader should be cautioned about possible outliers in these data. If there are outliers which should be removed, the block $\times$ treat-

TABLE 13.7. Variances and Comparisonwise Confidence Intervals for Effects in Table 13.4.

Variances and Confidence Intervals	Cultivar A	B	C	D	E	F
$\hat{V}(\hat{\delta}_{i(\cdot)})$	2.8373	0.0385	12.694	0.0576	1.1583	0.0352
$t_{.05,22}\sqrt{\hat{V}(\hat{\delta}_{i(\cdot)})}$	3.49	0.407	7.39	0.498	2.23	0.389
$\hat{V}(\hat{\delta}_{i(j)})$	5.5886	0.0758	25.003	0.1134	2.2815	0.0693
$t_{.05,22}\sqrt{\hat{V}(\hat{\delta}_{i(j)})}$	4.90	0.571	10.4	0.698	3.13	0.546
$\hat{V}(\hat{\pi}_{i(jg)})$	1.2897	0.0175	5.7700	0.0262	0.5265	0.0160
$t_{.05,22}\sqrt{\hat{V}(\hat{\pi}_{i(jg)})}$	2.36	0.274	4.98	0.335	1.50	0.262
$\hat{V}(\hat{\beta}_{i(all\ but\ i)})$	6.1904	0.0840	27.696	0.1256	2.5272	0.0768
$t_{.05,22}\sqrt{\hat{V}(\hat{\beta}_{i(all\ but\ i)})}$	5.16	0.601	10.91	0.735	3.30	0.575

ment mean squares would be decreased and more significant results would be obtained.

13.4 Combined Responses for Three or More Crops

As stated in Chapter 4 of Volume I, the grower of crops in a farming system would be interested in some linear combination of crop responses. In most cases, the responses will be weights of fruit, grain, fodder, biomass, or some other characteristic. In some cases, the response could involve numbers rather than weight, e.g., oranges, ears of sweet corn, etc. Whatever the response of interest, it will appear in the weighted total response for the system. For v crops in mixtures of size k, $k = 1, \ldots, v$, and responses Y_{hij}, the linear combination would be

$$\sum_{h=1}^{v} a_h Y_{hij} = Z_{ij}, \tag{13.43}$$

where a_h is a weighting factor for crop h in block j in the ith mixture of size k; all crops not appearing in the mixture will have $a_h = 0$. Thus, for a sole crop, all $a_h = 0$ except one, for a mixture of two crops, only two of the a_h will be nonzero. Equation (13.43) is a generalization of the results in Chapter 4 of Volume I (also, see Federer, 1987, Riley, 1984).

From an economic point of view, a_h is the value of crop h. From a nutritional point of view, a_h would represent a calorie or a protein conversion factor. From a land use point of view, a_h would be the reciprocal of the sole crop yield and Y_h would be the yield of crop h in the mixture. From a statistical viewpoint, (13.43) could be the linear combination maximizing the variance of the linear combination or which maximized the treatment sum of squares divided by the treatment plus error sums of squares. The statistical view does not lend itself to practical interpretation and, hence, would not ordinarily be of use in an experiment on intercropping. Since ratios of prices and ratios of sole crop yields are much more stable than are prices or yields themselves, it is recommended that one crop be selected as a base crop and that the coefficients for all crops be divided by this coefficient. Thus, if $h = 1$ is the base crop, then (13.43) becomes

$$\sum_{h=1}^{v} \frac{a_h Y_h}{a_1} = \sum_{h=1}^{v} b_h Y_h, \tag{13.44}$$

where $b_h = a_h/a_1$. When a_h is the reciprocal of the sole crop yields, say Y_{sh}, then (13.44) becomes

$$\text{LER}^* = \sum_{h=1}^{v} Y_{s1} Y_h / Y_{sh} = Y_{s1} \sum_{h=1}^{v} L_h, \tag{13.45}$$

where $L_h = Y_h/Y_{sh}$, Y_{sh} is the sole crop yield for crop h, Y_{s1} is the yield for the base sole crop, and Y_h is the yield of crop h in the mixture. Equations (13.44) and (13.45) would be called relative linear combinations, e.g., relative economic values, relative land use or land equivalent ratios, relative calories or protein, etc. For comparative purposes of cropping systems, such relative values are useful and, as shown in Chapter 4 of Volume I and Chapter 11, they can be discussed together simply by changing the values for a_h or b_h.

The values for Y_{sh} should be fixed values and not random variables. If Y_h and Y_{sh} have a bivariate normal distribution, then the variance of $Y_1/Y_{s1} + Y_2/Y_{s2}$ is infinite (Federer and Schwager, 1982). Of course, the assumption of normality for Y_{sh} or Y_h is invalid since yields cannot be negative. Perhaps some bivariate distribution, with non-negative values for Y_h and Y_{sh}, e.g., gamma, chi square, etc., would yield a distribution for (13.45) with a finite variance. This problem requires investigation. However, if Y_{sh} or Y_{sh}/Y_{s1} are constants, the problem of infinite variance does not arise. It is recommended that Y_{sh} be the average yield from growers' fields over a period of years in the region for which the proposed cropping systems are to be used. Mead and Willey (1980) and Mead and Riley (1981) have recommended use of "optimal" yield. Since some estimate of average growers' yields would be available for most regions and since the concept of optimal yield would usually be hard to define and may vary with individuals, it is recommended that the average ratio of growers' yields be used. Then, b_h can be regarded as a constant, and the problem of infinite variances does not arise. It would appear that such a procedure would be more in line with what a grower would do.

Regardless of the linear combination used to combine values across all crops, the statistical analysis procedures are for a standard univariate analysis. The standard randomized complete blocks analysis of variance, F-tests, multiple comparisons procedures, etc. may be used. There should be no trouble with variance heterogeneity, or perhaps even the assumption of normality. All analyses will be conditional upon the ratios used. As in Chapter 4 of Volume I, it is recommended that a range of ratios be used; for example, a low, medium, and high value. For v crops, this would require 3^{v-1} ratio combinations. For 3 crops, one could construct 9 graphs like Figure 4.1 of Volume I; for four crops, 27 such graphs would be required; etc. For such quantities as total calories, starch, or protein, many fewer ratio combinations would probably be required. The experimenter should give considerable thought to this prior to doing the computations.

The $c_1c_2c_3$ mixtures could be treated as a three-factor factorial with c_1 levels of factor (crop) one, c_2 levels of factor (crop) two, and c_3 levels of factor (crop) three. The lines form the levels of crops and are unordered. For mixtures of v crops, one could consider a v-factor factorial with $c_1c_2 \cdots c_v$ combinations. Any of the functions of the v responses for a mixture of lines of v crops could be treated in this fashion.

The previous discussion deals with linear combinations of data according to some specific plan. One could let the data determine which linear combination(s) would most effectively discriminate among the treatments (lines of the crops) by using a multivariate analysis approach. The v crops would be considered to be v varieties in the MANOVA (multivariate analysis of variance). A table similar to Table 4.5 of Volume I could be used, where the matrix would be $v \times v$ instead of 2×2. If mixtures of k of v crops were used, the article by Srivastava (1968) may be helpful in using a multivariate analysis.

Given that Z_{ij} values are available from an experiment laid out as a randomized complete block design, e.g., where Z_{ij} is the yield of the ith system (treatment) from the jth block, standard analyses of variance, multiple comparisons procedures, significance tests, interval estimation, etc. may be conducted. Of course, as usual, variance homoscedasticity, additivity, and independence among the Z_{ij} must hold for most procedures. To illustrate the use of created variables such as those in (13.43) to (13.45), the data from Examples 13.3 and 13.5 plus a third example are utilized.

Example 13.6. An experiment involving three cultivars (cotton, maize, and beans) in all combinations of one, two, and three crops was conducted in 1975 and 1976 at the Barbalha Experiment Station in Ceará, Brasil. The yields of the three sole crops, three mixtures of two cultivars, and one mixture of all three crops are given in Table 13.8. If the monetary worth of cotton is five times that of maize and the monetary value of beans is four times that of maize, then forming the linear combinations $Z_{ij} = 5C_{ij} + M_{ij} + 4B_{ij}$, where C_{ij} is the yield of cotton, M_{ij} is the yield of maize, and B_{ij} is the yield of beans for mixture i in year j, $i = 1, 2, \ldots, 7$ treatments and $j = 1975, 1976$. For example, the created relative monetary value for $i = 6$ or treatment M + B in 1975 is $4610 = 5(0) + 2406 + 4(551)$. An analysis of variance on this created variable is given at the bottom of Table 13.8, along with an F-statistic for treatments. It should be noted (i) a multiple comparisons procedure or (ii) a 2^3-factorial analysis for a fractional replicate could have been used. (Note: The no C, no M, and no B combination produces zero yield, but this combination, while completing the 2^3 factorial, should not be included in the

TABLE 13.8. Experiment on Mixtures Conducted at the Barbalha Experiment Station, Ceará, Brazil. Production of Cotton, Maize, and Beans in kg/ha, 1975 and 1976.

	1975			1976		
Treatments	C	M	B	C	M	B
Cotton = C	408	—	—	240	—	—
Maize = M	—	3102	—	—	3443	—
Beans = B	—	—	1007	—	—	1316
C + M	140	294	—	76	3269	—
C + B	366	—	826	308	—	1139
M + B	—	2406	551	—	2771	555
C + M + B	174	2135	515	161	2368	531

	Monetary $5C_{ij} + M_{ij} + 4B_{ij}$				$C_{ij} + (\bar{c}/\bar{m})M_{ij} + (\bar{c}/\bar{b})B_{ij}$ = LER*			
Treatments	1975	1976	Total	Mean	1975	1976	Total	Mean
Cotton = C	2040	1200	3240	1620.0	408	240	648	324.0
Maize = M	3102	3443	6545	3272.5	310	344	654	327.0
Beans = B	4028	5264	9292	4636.0	282	368	650	325.0
C + M	3641	3649	7290	3645.0	434	403	837	418.5
C + B	5134	6096	11,230	5615.0	597	627	1224	612.0
M + B	4610	4991	9601	4800.5	395	432	827	413.5
C + M + B	5065	5297	10,362	5181.0	532	546	1078	539.0
Total	27,620	29,940	57,560	4111.4	2958	2960	5918	422.7

		Monetary			Production = LER*$/\bar{c}$		
Source of variation	d.f.	Sum of squares	Mean square	F	Sum of squares	Mean square	F
Total	14	260,978,602			2,677,540		
CFM	1	236,653,829			2,501,623		
Year	1	384,457			0		
Treatment	6	22,587,738	3,764,623	16.7	155,816	25,969	7.8
Year × treatment	6	1,352,578	225,430		20,101	3,350	

analysis.) It is suggested that (i) is the appropriate procedure to use, as one wants the best combination, and it is left as an exercise for the reader.

The investigator might wish to consider land use or relative land use rather than relative monetary values. For this, let cotton rather than maize be the standard crop. Then, the coefficient of C_{ij} is unity, of M_{ij} it is $(408 + 240)/(3102 + 3443) = 0.10 = \bar{c}/\bar{m}$, and of B_{ij} it is $\bar{c}/\bar{b} = (408 + 240)/(1007 + 1316) = 0.28$. The variable called production in Table 13.8 is obtained by taking the value for the preceding variable, denoted as LER*, and divide it by $\bar{c}$ to obtain a LER (land equivalent ratio). An analysis of variance on relative LERs is given at the bottom of Table 13.8, together with an F-statistic. The analysis is valid given that $\bar{c}/\bar{m}$ and $\bar{c}/\bar{b}$ are constants and not random variables. Here, again, it is suggested that a multiple comparisons procedure would be more appropriate than an F-statistic, as the experimenter usually will want the mixture giving the largest value.

The results for the seven treatments are depicted in Figure 13.1. The production (LER) and the relative land equivalent ratio, LER*, must give the same pattern in the graph, as they differ only by a multiple of $\bar{c}$. The axes could be adjusted so that both plots would be the same. The monetary pattern is different from the LER and LER* because the coefficients are different. The investigator will need to insert appropriate values of the b_h in (13.44) for each particular situation. If protein, carbohydrate, or oil values were of interest, the b_h would be different from each other and from the above. The patterns would also differ from those in Figure 13.1 because different linear combinations of the yields are being obtained.

Example 13.7. The data of Example 13.3 are used here. Maize is considered to be the base crop and the following created variables were obtained:

$$S_{ij} + 3C_{ij} + 3B_{ij} + M_{ij} = Z_{1ij},$$

$$S_{ij} + 5C_{ij} + 5B_{ij} + M_{ij} = Z_{2ij},$$

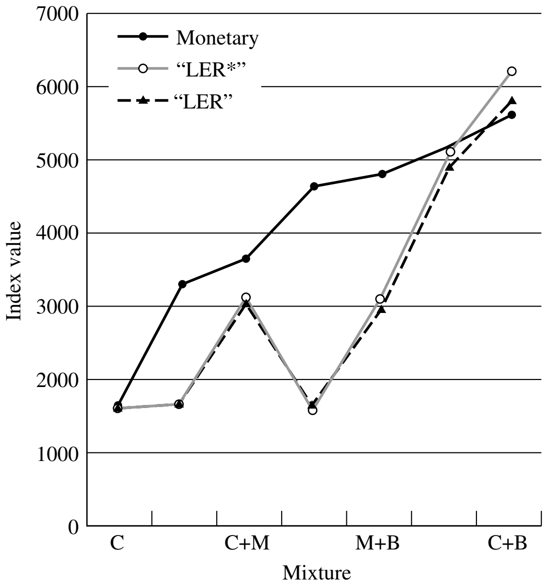

FIGURE 13.1. Monetary data from Table 13.8. "LER*" is LER* minus 224 multiplied by 16. "LER" is LER minus 2/3 multiplied by 4800. (These transformations were made in order to put all three indices on the same graph.)

and

$$0.9S_{ij} + 8C_{ij} + 7B_{ij} + M_{ij} = Z_{3ij},$$

where C_{ij}, B_{ij}, and M_{ij} are as in Example 13.6 and S_{ij} is the yield of sorghum in mixture i and location j = I, II. The values of Z_{1ij}, Z_{2ij}, and Z_{3ij} are given in Table 13.9. Analyses of variance on these created variables are presented at the bottom of Table 13.9. Here, again, a multiple comparisons procedure would be the preferred analysis for most investigators.

An investigator may wish to consider land use by using (13.45) or relative land equivalent ratios. If so and if the relative yields of crops in the region of inference is a reflection of those in the experiment, then the created variable totals

$$\begin{aligned} Z_{4i\cdot} &= M_{i\cdot} + \frac{\bar{m}}{\bar{c}} C_{i\cdot} + \frac{\bar{m}}{\bar{s}} S_{i\cdot} + \frac{\bar{m}}{\bar{b}} B_{i\cdot} \\ &= M_{i\cdot} + 2.5C_{i\cdot} + S_{i\cdot} + 5B_{i\cdot} \end{aligned}$$

are 5265, 5129, 5279, 6793, 7601, 5410, 6580, and 8205, respectively, for treatments B, S, M, M+B, S+B, C, M+C, and S+C. These values are identical to those for $Z_{2i\cdot}$ for the first five treatments but differ for the last three. Depending on which variable is of interest to the grower or the investigator, the cropping system or treatment which is optimal varies with the variable. C is best for Z_3; S+C is best for Z_1, Z_2, and Z_4.

A study of the original data or of the created variables Z_{fij}, $f = 1, 2, 3, 4$, indicate that heterogeneity of error variance is present. The erratic responses of beans alone or in mixtures is the cause of this heterogeneity. The error variance denoted as R×T (elim. B), which is the year by treatment without beans, is considerably lower than those involving treatments with beans present. For Z_1, the ratios of variance are 1:2.5:17; for Z_2, the ratios are 1:3.9:26; and for Z_3, the ratios are 1:4:28. Since heterogeneity of error variances is present, it is suggested that the Behrens or generalized Behrens procedure be used (see Grimes and Federer, 1984).

In Figure 13.2, the total values $Z_{1i\cdot}$, $Z_{2i\cdot}$, $Z_{3i\cdot}$, and $Z_{4i\cdot}$ are plotted against treatments ordered by variable Z_1. If the variable Z_4 is the one of interest to an experimenter, then treatments S+C and S+B are the best. If variable Z_3 were the one of interest, then the sole crop cotton and the mixture S+C were the two highest. When the variable of interest is selected, the results may be graphed as in Figure 13.2. In this manner, the investigator may compare different cropping systems as well as having an idea of variation expected for the particular cropping system.

Example 13.8. For the data from Example 13.5, we may create variables, as has been done for the previous two examples. Here, however, we use a negative weight for one of the cultivars in one of the created variables. This is for A wild oat and variable Z_1 . Since wild oat would not normally have any monetary value nor food value for humans, it may be costly to remove this crop from a mixture. It could also have a beneficial value for yields of other crops in a mixture, but could have zero monetary or food value. If it were beneficial in a mixture, wild oat may be

TABLE 13.9. Created Variables for Various Coefficients of Crop Yields for Data of Example 13.3. (SS = Sum of Squares, MS = Mean Squares.)

Treatment	$Z_{1ij} = S + 3C + 3B + M$ I	II	Total $Z_{1i\cdot}$	$Z_{2ij} = S + 5C + 5B + M$ I	II	Total $Z_{2i\cdot}$	$Z_{3ij} = 0.9S + 8C + 7B + M$ I	II	Total $Z_{3i\cdot}$
S	2210	2919	5129	2210	2919	5129	1989	2627	4616
C	3660	2832	6492	6100	4720	10820	9760	7552	17312
B	207	2952	3159	345	4920	5265	483	6888	7371
M	2451	2828	5279	2451	2828	5279	2451	2828	5279
S + C	4440	4338	8778	5682	5386	11068	7287	6681	13968
S + B	2293	4066	6359	2359	5042	7401	2206	5758	7964
M + C	3645	3277	6922	4483	3775	8258	5740	4522	10262
M + B	2689	3530	6219	2749	4044	6793	2809	4558	7367
Total	21,595	26,742	48,337	26,379	33,634	60,013	32,725	41,414	74,139

Analyses of Variance for Created Variables

Source of variation	d.f.	Z_{1ij} SS	MS	Z_{2ij} SS	MS	Z_{3ij} SS	MS
Total	16	161,621,047	—	261,763,087	—	442,649,431	—
Correction for mean	1	146,029,098	—	225,097,511	—	343,536,958	—
Year = R	1	1,655,726	—	3,289,689	—	4,718,670	—
Treatment = T	7	9,160,920	1,308,703	20,193,462	2,884,780	67,125,010	9,589,287
R × T	7	4,775,303	682,186	13,182,425	1,883,204	27,268,793	3,895,542
R × T (elim. B)	4	733,617	183,404	1,400,565	350,141	2,727,370	681,842
R × T (for treat. w/B)	2	906,437	453,219	2,710,768	1,355,384	5,511,459	2,755,730
R × (Treat. w/B vs. Treat. w/o B)	1	3,135,249	—	9,071,093	—	19,029,965	—

included in the mixture, but it would have a weight of zero in a created variable for this situation. For Example 13.5, A is wild oat, B is coriander, C is lentils, D is birdsfoot trefoil, E is alfalfa, and F is chamomile. Suppose the following variable

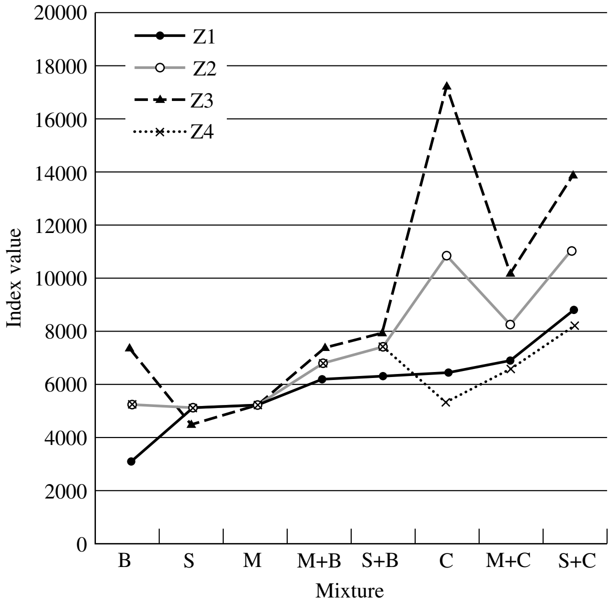

FIGURE 13.2. Total responses for created variables Z_1 to Z_4 plotted against cropping system ordered on Z_1.

is created for treatment i in block j:

$$Z_{1ij} = -0.5A_{ij} + 10B_{ij} + 5C_{ij} + D_{ij} + E_{ij} + 10F_{ij}.$$

Here, alfalfa, E_{ij}, is set as the base crop with birdsfoot trefoil, D_{ij}, having equal value or use. Lentils might be considered as 5 times as valuable as alfalfa, and coriander and chamomile could have a value of 10 times alfalfa on a weight basis. Wild oat is given a negative weight which would be appropriate if it was to be removed from a mixture, i.e., it would be considered as a weed with only a detrimental effect if it is not removed. Also, note that a cultivar could be costly to include, but it may have a beneficial effect on other crops in a mixture. In this case, a negative coefficient may or may not be used.

Another created variable might be land equivalent ratios or relative land equivalent ratios. In absence of other information, let us use the ratio of yields in this experiment as representative of the region where these mixtures would be grown. Taking C, lentil, as the base crop, then the variable Z_2, relative land equivalent ratio from (13.45), is created as

$$Z_{2ij} = 2A_{ij} + 24B_{ij} + C_{ij} + 15D_{ij} + 4E_{ij} + 25F_{ij};$$

where $118.63/67.75 \doteq 2, 118.63/4.91 \doteq 24, 118.63/8.15 \doteq 15, 118.63/26.90 \doteq 4$, and $118.63/4.78 \doteq 25$ (see Table 13.2). Also, let

$$Z_{3ij} = A_{ij} + B_{ij} + C_{ij} + D_{ij} + E_{ij} + F_{ij}$$

be the total biomass. Note that if a crop is not present, its yield is zero.

TABLE 13.10. Created Variables from Data in Table 13.2 for Variables Z_1, Z_2, and Z_3.

Treatment	Variable			Treatment	Variable		
i	$Z_{1i\cdot}$	$Z_{1i\cdot}$	$Z_{1i\cdot}$	i	$Z_{1i\cdot}$	$Z_{1i\cdot}$	$Z_{1i\cdot}$
A	−34	136	68	ADF	8	114	24
B	49	118	5	AEF	22	159	42
C	593	119	119	BCD	194	119	37
D	8	122	8	BCE	353	194	74
E	27	108	27	BCF	321	177	59
F	48	120	5	BDE	40	140	16
ABC	277	130	68	BDF	21	57	3
ABD	7	135	32	BEF	42	115	12
ABE	30	152	33	CDE	262	120	62
ABF	10	116	31	CDF	359	147	72
ACD	216	116	63	CEF	301	145	63
ACE	224	108	73	DEF	25	98	12
ACF	280	162	78	ABCDEF	216	150	55
ADE	−2	96	30	—	—	—	—

Variables are

$$Z_{1i\cdot} = -0.5A_{i\cdot} + 10B_{i\cdot} + 5C_{i\cdot} + D_{i\cdot} + E_{i\cdot} + 10F_{i\cdot},$$

$$Z_{2i\cdot} = 2A_{i\cdot} + 24B_{i\cdot} + C_{i\cdot} + 15D_{i\cdot} + 4E_{i\cdot} + 25F_{i\cdot},$$

$$Z_{3i\cdot} = A_{i\cdot} + B_{i\cdot} + C_{i\cdot} + D_{i\cdot} + E_{i\cdot} + F_{i\cdot},$$

Using the totals from Table 13.2, the various totals $Z_{1i\cdot}$, $Z_{2i\cdot}$, and $Z_{3i\cdot}$ may be computed and are presented in Table 13.10. These results are depicted graphically in Figure 13.3. For the variable Z_1, the relatively high yield of C together with a coefficient of 5 separate the 27 treatments into two groups, those with C and those without C. For variable Z_2, there is no separation, but there is again for variable Z_3 even if not so drastic. The two mixtures BCE and BCF have the highest land use values, whereas C and ACF have the highest biomass yield.

For crops such as these, it would appear that the calculation of land use values, land equivalent ratios, would not be what an investigator would desire. There are cases in intercropping where land equivalent ratios are inappropriate and this experiment would be one of them. This means that investigators who use land equivalent ratios for *all* intercropping investigations could be doing inappropriate analyses for some of their experiments. Here, again, we demonstrate that *the statistical analysis of an intercropping investigation must be for the goals, nature, and limitations of the particular experiment being analyzed*. It is inappropriate to use only one analysis for all investigations.

Another aspect of a statistical analysis is to compare a variable Z_f, $f = 1, 2, 3$, with the yields of the sole crops for the same mixture. Consider variable Z_3. For the mixture ABC, the average of the sole crop yields on a four-plant basis, one-third of a plot, is $(67.55 + 4.91 + 118.63)/3 = 64 = Z_{5ABC}$, which is the standard variable with which to compare Z_{3ABC}. The first $Z_{3i\cdot}$ value is computed as follows

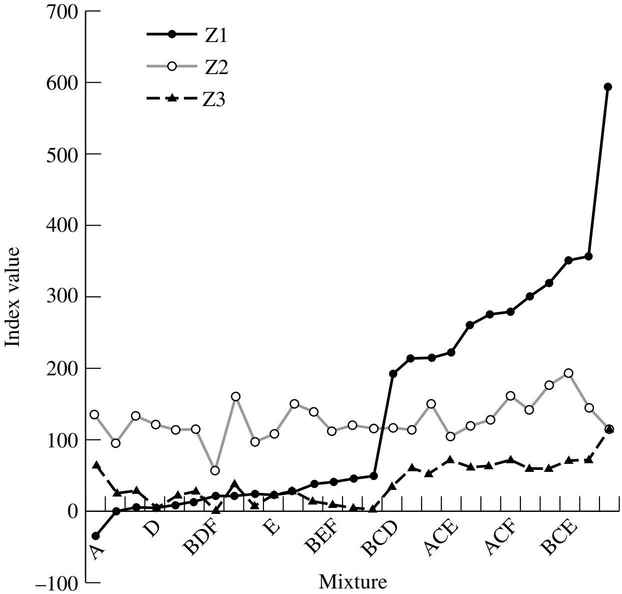

FIGURE 13.3. Responses for Z_1, Z_2, and Z_3 versus treatments ordered on Z_1. (Treatment order is A, ADE, ABD, D, ADF, ABF, BDF, AEF, DEF, E, ABE, BDE, BEF, F, B, BCD, ACD, ABCDEF, ACE, CDE, ABC, ACF, CEF, BCF, BCE, CDF, C).

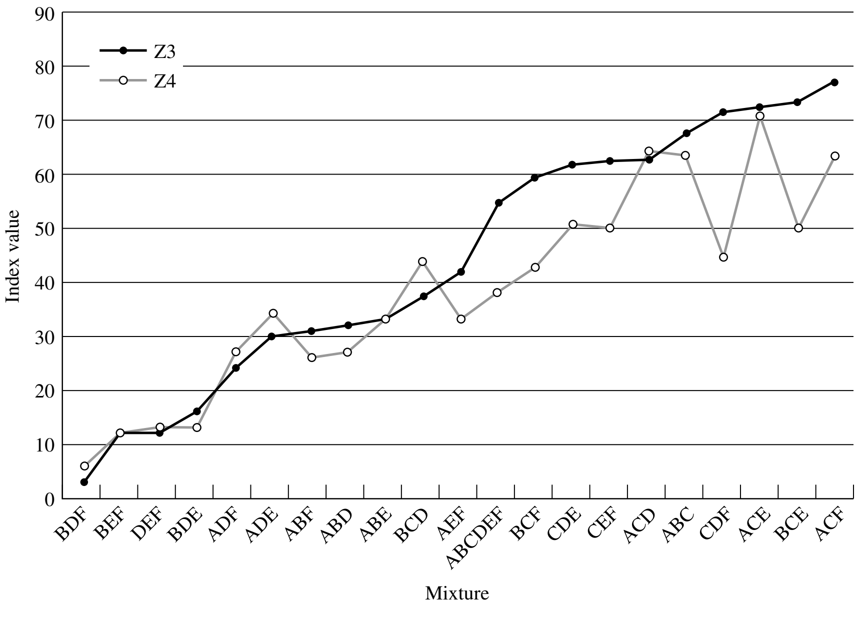

FIGURE 13.4. Responses for Z_3 and Z_5 plotted against mixtures of three crops ordered by Z_3.

using the yields from Table 13.2:

$$Z_{3ABC} = 13.88 + 2.08 + 52.54 = 68.$$

The remaining $Z_{3i\cdot}$ and $Z_{4i\cdot}$ values are computed in a similar manner and are given in Table 13.11 and depicted graphically in Figure 13.4. From the graph, it may be noted that as the biomass increases, the deviations $Z_{3i\cdot} - Z_{4i\cdot}$ increase, and that the crop C is involved. The question arises as to whether or not the mean of the deviations $Z_{3i\cdot} - Z_{4i\cdot}$ differs significantly from zero. The null hypothesis is that the mean of $Z_{3i\cdot} - Z_{4i\cdot}$ is zero. To test this, the mean of the deviations is $[4+5+0+\cdots+(-1)+17]/21 = 131/21 = 6.24$, the variance of this mean is $[4^2+5^2+0^2+\cdots+(-1)^2+17^2-131^2/21]/20(21) = 4.3376$, and the t-statistic is $t = 6.24/\sqrt{4.3376} = 3.00$. The tabulated $t_{.02,20} = 2.52$ and $t_{.01,20} = 2.84$. Alternatively, a chi-square test could have been performed using the number of plus deviations and the number of negative deviations. More significant figures could have been used to eliminate the zero deviations. Otherwise, the number of zeros could be equally divided between the pluses and minuses. Thus, χ^2(1 d.f.) $= [(14-(21/2))^2+(7-(21/2))^2]/21/2 = 2.33$. This test does not take into account the size of the differences as does the t-statistic. The negative differences are small, whereas the positive ones are relatively large.

The idea used to compare $Z_{3i\cdot}$ and $Z_{4i\cdot}$ responses may be carried further. We may look at the yields of cultivars in mixtures and compare these with sole crop yields after making the totals comparable. For example, consider a single crop A. The yield of A in all the mixtures of three where it appeared is $3(13.88+27.75+20.50+28.47+15.81+21.82+22.46+21.16+20.68+24.03)/10 = 65$, where the coefficient of 3 is used to put yields on the same basis as the sole crop yields (A in the mixture occupied only one-third of an experimental unit) and 10 is the number of mixtures of three where A occurred. This value of 65 may be compared with the sole crop yield $67.75 \doteq 68$. The remaining values are computed in a similar manner and are given in Table 13.12. The striking item here is the large deviation, 36, for lentils, C. The sum of the remaining deviations is near zero. If lentils, C, were the crop of interest, then mixtures would make better use of the land.

An investigator may wish to observe what happens with pairs of crops in a mixture of three. For example, the yield for crops A and B in the four mixtures in which they appear (see Table 13.2) is $3(13.88+2.08+27.75+1.81+20.50+3.04+28.47+1.67)/4 = 74$, where 3 is needed to put these yields on the same basis as the sole crop. The sum of the yields of sole crops A and B is $67.75+4.91 = 73$. The remaining yields of pairs and sole crops are computed in the same manner and are given in Table 13.12. The deviations are also computed. Any pair where crop C occurs produces a large positive deviation indicating how well C does in mixtures. Here, again, we may compute the mean and standard error of these deviations (the sum of the 15 deviations is 167) and compute a t-statistic as $t = (167/15)/\sqrt{4659.73/14(15)} = 11.13/4.71 = 2.36$, $t_{.05,14} = 2.15$, and $t_{.02,14} = 2.62$. A significant difference from zero is indicated at the 5% level.

TABLE 13.11. Variable Z_3 Compared with Mean Yields from Sole Crops. ($Z_{4i\cdot}$ = Mean of Sole Crop Yields for Crops in Treatment i.)

Treatment	Variable		
i	$Z_{3i\cdot}$	$Z_{4i\cdot}$	$Z_{3i\cdot} - Z_{4i\cdot}$
ABC	68	64	4
ABD	32	27	5
ABE	33	33	0
ABF	31	26	5
ACD	63	65	−2
ACE	73	71	2
ACF	78	64	14
ADE	30	34	−4
ADF	24	27	−3
AEF	42	33	9
BCD	37	44	−7
BCE	74	50	24
BCF	59	43	16
BDE	16	13	3
BDF	3	6	−3
BEF	12	12	0
CDE	62	51	11
CDF	72	44	28
CEF	63	50	13
DEF	12	13	−1
ABCDEF	55	38	17

If it is desired to compare each of the differences $Z_{3i\cdot} - Z_{4i\cdot}$ with zero, a range or t-test could be used for either a significance test or for computing confidence intervals. For example, consider treatment ABC. The variance of the difference $Z_{3ABC} - Z_{4ABC}$ is approximated by

$$V\left[\left(Y_{\cdot A(BC)} - Y_{\cdot A}/3\right) + \left(Y_{\cdot B(AC)} - Y_{\cdot B}/3\right) + \left(Y_{\cdot C(AB)} - Y_{\cdot C}/3\right)\right],$$

which is approximated by $2r\left(\hat{\sigma}^2_{\epsilon A} + \hat{\sigma}^2_{\epsilon B} + \hat{\sigma}^2_{\epsilon C}\right) = 2(3)(7.738+0.105+34.62) = 254.8$ (see Table 13.5). The assumption made here is that the variance of $Y_{\cdot A(BC)}$ and $Y_{\cdot A}/3$ are approximately the same for reasons described earlier in this chapter. The other variances may be computed in a similar manner. A $(1-\alpha)\% = 95\%$ confidence interval for $Z_{3ABC\cdot} - Z_{4ABC\cdot} = 68 - 64$ would be constructed as $(68-64) \pm t_{.05,22}\sqrt{254.8} = 4 \pm 33 = 37$ to -29. Before constructing the remaining confidence intervals and interpreting the results, it is again suggested that a study be made of the residuals with the idea that outlying observations may be present in this experiment.

Analyses of variance may be computed using the created variables as was done for Example 13.7 in Table 13.9. To illustrate, an analysis of variance for the variable Z_1 is

TABLE 13.12. Yields by Pairs of Crops and by Individual Crops in a Mixture of Three and Comparable Sole Crops Responses. (Mixture Yields are Multiplied by Three to Put Pixture Total on a 12-Plant Basis.)

	Yields				Yields		
Pairs	Mixture	Sole	Diff.	Crop	Mixture	Sole	Diff.
AB·	74	73	1	A··	65	68	−3
AC·	202	186	16	B··	7	5	2
AD·	71	76	−5	C··	155	119	36
AE·	94	95	−1	D··	6	8	−2
AF·	77	73	4	E··	28	27	1
BC·	159	124	35	F··	5	5	0
BD·	11	13	−2				
BE·	38	32	6				
BF·	10	10	0				
CD·	153	127	26				
CE·	181	146	35				
CF·	178	123	55				
DE·	33	35	−2				
DF·	9	13	−4				
EF·	35	32	3				

Source of variation	d.f.	SS	MS
Total—CFM	80	244,664	—
Block	2	2,172	1,086
Treatment	26	217,823	8,378
Block×treatment	52	24,669	474

All 6 single crops, 20 three-crop mixtures, and 1 six-crop mixture make up the 27 treatments. The ratio of the treatment mean square to block × treatment mean square is 17.7, whereas $F_{.05}(26, 52) = 1.71$. As may be observed from Figure 13.3, there are large differences among the 27 treatment totals. A large part of the treatment mean square is attributable to the single degree of freedom contrast of treatments with C present versus treatments with C absent. There are, however, large differences among the treatments containing C, but relatively small differences among those treatments where C was absent. The partitioning of the treatment sum of squares is left as an exercise for the reader. If desired, similar analyses of variance and F-statistics may be computed for other variables or even for differences of the form $Z_{3ij} - Z_{4ij}$. For the latter, the sole crops would need to be deleted and an analysis of variance computed for the remaining 21 treatment differences from the three blocks. Partitioning of treatment degrees of freedom or multiple comparisons procedures may elicit the information desired by the investigator.

13.5 Some Comments

The examples presented demonstrate the diversity in types of intercropping experiments and in types of statistical analyses that are useful in eliciting the information contained in the experiments. Sole cropping ideas and goals need to be extended considerably in order to provide appropriate analyses and interpretations for intercropping experiments. From Chapter 11, we reiterate that the investigator should "expect the unexpected" from an intercropping experiment. Several results which were unusual and unexpected to the writer occurred in the examples here, just as they did in the preceding chapter. The large difference in bean yields at the two locations in Example 13.3 is striking. Such differences appear to be extraordinarily large. The investigator should provide an explanation as to why the differences were so large. Also, bean and cowpea yields were reduced more when grown with maize than with sorghum. Is this a varietal or species phenomenon? Why?

In Example 13.4, cotton yields were relatively unaffected when fertilized. This would mean that any additional yield from the intercrops is obtained as an additional bonus. Why would cotton yields be unaffected when intercropped in this manner? What is the nature and physiology of cotton which allows this to happen? Why would maize yields be lower (or the same) on fertilized than on unfertilized plots?

Example 13.5 had considerable variation among the experimental units treated alike. The large coefficients of variation indicate that experimental technique needs to be reconsidered. The size of the experimental unit (see Figure 13.1) immediately comes to mind.

Many of the computations described in this chapter are programmable using such software packages as SAS or GENSTAT. It is suggested that the computations be done on a pocket or desk calculator until the analyst becomes familiar with the statistical models and analyses. Then, the packages may be used for calculations.

13.6 Problems

13.1 For crop D, recompute the analysis using 0.27 rather than 0.22 (Table 13.2) for combination ADF in block three.

13.2 Obtain the analyses described in this chapter after making the transformation log(yield+1) for the data in Table 13.2. Note that 1 is added to the yield of cultivars D and F since some of the values are near zero. Do this for all cultivars.

13.3 For the data of Table 13.2 corrected as in Problem 1, compute the residuals and determine if there are possible outliers. If so, one would need to question the experimenter as to possible reasons for this.

13.4 Select a multiple comparisons procedure and make the appropriate comparisons and interpretations for the data of Example 13.6.

13.5 Partition the treatment sum of squares with 26 degrees of freedom as suggested in the text following Table 13.12. Do likewise with the error sum of squares. Make the appropriate interpretations.

13.6 For the data following equation (13.48), verify that equations (13.49) through (13.57) are correct by performing the calculations.

13.7 Verify that equations (13.61) to (13.70) hold for the parameter values in the text following equations (13.58) – (13.60). Verify the totals following equation (13.70).

13.8 Given a canonical variate $Y_1 + bY_2 + cY_3$, show how to extend the computer program in Chapter 4 of Volume I to obtain values of b and c which maximize (treatment sums of squares)/(treatment + error sum of squares). How would you extend this to include four variables?

13.9 What effect does removing the value of 207 for treatment B for Z_{1ij} have on the analysis of variance and on F-tests? Are there more outliers for these data?

13.7 Literature Cited

Aidar, H. (1978). Estudo sobre populacoes de plants emdois de culturas associades de milho e feijao. Viscosa, M.G., Tese de Doutorada, Universidade Federalde Vicosa.

Federer, W.T. (1987). Statistical analyses for intercropping experiments. *Proc., Thirty-Second Conference on the Design of Experiments in Army Research Development and Testing*, ARO Report 87-2, pp. 1-29.

Federer, W.T. and S.J. Schwager (1982). On the distribution of land equivalent ratios. BU-777-M in the Technical Report Series of the Biometrics Unit, Cornell University, Ithaca, NY.

Grimes, B.A. and W.T. Federer (1984). Comparison of means from populations with unequal variances. In *W.G. Cochran's Impact on Statistics* (Editors: P.S.R.S. Rao and J. Sedransk), John Wiley & Sons, Inc., New York, pp. 353-374.

Mead, R. and J. Riley (1981). A review of statistical ideas relevant to intercropping research (with discussion). *J. Roy. Statist. Soc., Ser. A* **144**, 462-509.

Mead, R. and R. Willey (1980). The concept of 'Land equivalent ratio' and advantages in yields from intercropping. *Exp. Agric.* **16**, 217-228.

Riley, J. (1984). A general form of the 'Land Equivalent Ratio.' *Exp. Agric.* **20**, 19-29.

Srivastava, J.N. (1968). On a general class of designs for multiresponse experiments. *Ann. Math. Statist.* **39**, 1825-1843.

Appendix 13.1

Response Equations for a 12-Plant Basis

For the experiment and data of Example 13.5, let us consider the response equations to be (for cultivar A)

Sole crop for A (yield from 12 plants):

$$Y_{hA} = \mu + \tau_A + \rho_h + \epsilon_{hA}. \tag{13.46}$$

Mixture of three crops, A with j and g ≠ A (yields from four plants):

$$Y_{hA(jg)} = \frac{1}{3}(\mu + \tau_A + \rho_{hA}) + \frac{2}{3}\left(\delta_{A(j)} + \delta_{A(g)}\right) + \pi_{A(jg)} + \epsilon_{hA(jg)}. \tag{13.47}$$

Mixture of six crops, A with five others (yields from two plants):

$$Y_{hA(BCDEF)} = \frac{1}{6}(\mu + \tau_A + \rho_{hA}) + \frac{1}{3}\sum_{j=B}^{F}\delta_{A(j)} + \beta_{A(BCDEF)} + \epsilon_{hA(BCDEF)}. \tag{13.48}$$

The coefficient of 1/3 in (13.47) is to put $\mu + \tau_A + \rho_{hA}$ on the same basis as in (13.46). The coefficient of 2/3 in (13.47) is used because $\delta_{A(j)}$ should have been derived from 8 instead of 12 plants in order to be on the same basis as $\mu + \tau_A$. A similar explanation holds for the coefficients in (13.48).

For the above response equations, a set of normal equations after applying the parameter constraints

$$0 = \sum_{h=1}^{r}\rho_h = \sum_{\substack{j=1\\ \neq A,g}}^{v}\pi_{A(jg)} = \sum_{\substack{g=1\\ \neq A,j}}^{v}\pi_{A(jg)}, \sum_{\substack{j=1\\ \neq A}}^{v}\delta_{A(j\cdot)} = (v-1)\bar{\delta}_{A(\cdot)},$$

and $v = 6$ is

$$\sum_h Y_{hA} = Y_{\cdot A} = r(\mu + \tau_A), \tag{13.49}$$

$$\sum_h \sum_{g \neq j,A} Y_{hA(jg)} = Y_{\cdot A(j\cdot)} = \frac{r(v-2)}{3}(\mu + \tau_A) + \frac{2r(v-3)}{3}\delta_{A(j)} + \frac{2r(v-1)}{3}\bar{\delta}_{A(\cdot)}, \tag{13.50}$$

$$\sum_h \sum_j \sum_g Y_{hA(jg)} = Y_{\cdot A(\cdot\cdot)} = \frac{r(v-1)(v-2)}{2(3)}\left(\mu + \tau_A + 4\bar{\delta}_{A(\cdot)}\right), \tag{13.51}$$

$$\sum_{h=1}^{r} Y_{hA(BCDEF)} = Y_{\cdot A(BCDEF)}$$

$$= \frac{r}{6}(\mu + \tau_A) + \frac{r(v-1)}{3}\bar{\delta}_A + r\beta_{A(BCDEF)}. \tag{13.52}$$

In the above, $\sum_{j \neq A} \delta_{A(j)} = (v-1)\bar{\delta}_{A(\cdot)}$ and $v = 6$. v instead of 6 was used in the above equations to illustrate how extensions could be made for other values of v. The term $\delta_{A(j')}$ is equal to $\delta_A + \gamma_{A(j)}$ in equations (13.4)–(13.6) where $\gamma_{A(j)}$ is an interaction effect. Note that the error terms in (13.46) to (13.48), i.e., ϵ_{hA}, $\epsilon_{hA(jg)}$, and $\epsilon_{hA(BCDEF)}$ all have the same expected value zero but different variances. Since there was only one observation per block for the sole crop and for a mixture of all six crops, estimates of corresponding variances are not obtainable from a blocked experiment. The moment solutions for the parameters from equations (13.49) to (13.52) are

$$\hat{\mu} + \hat{\tau}_A = Y_{\cdot A/r} = \bar{y}_{\cdot A}, \tag{13.53}$$

$$\hat{\delta}_{A(\cdot)} = \sum_{j \neq A} \hat{\delta}_{A(j)}/(v-1)$$

$$= \frac{3}{2r(v-1)(v-2)} Y_{\cdot A(\cdot\cdot)} - \bar{y}_{\cdot A}/4, \tag{13.54}$$

$$\hat{\delta}_{A(j)} = Y_{\cdot A(j\cdot)}\left(\frac{3}{2r(v-3)}\right)$$

$$- Y_{\cdot A(\cdot\cdot)}\left(\frac{3}{2r(v-2)(v-3)}\right) - \frac{\bar{y}_{\cdot A}}{4}, \tag{13.55}$$

$$\hat{\pi}_{A(jg)} = \frac{1}{r} Y_{\cdot A(jg)} - \frac{1}{r(v-3)}\left(Y_{\cdot A(\cdot j)} + Y_{\cdot A(g\cdot)}\right)$$

$$+ \frac{2}{r(v-2)(v-3)} Y_{\cdot A(\cdot\cdot)}, \tag{13.56}$$

$$\hat{\beta}_{A(BCDEF)} = \frac{1}{r} Y_{\cdot A(BCDEF)} - \frac{1}{2r(v-2)} Y_{\cdot A(\cdot\cdot)} + \frac{(v-3)}{12}\bar{y}_{\cdot A}. \tag{13.57}$$

As a small numerical example to illustrate the above equations, consider the following set of parameters used to construct the observations and totals from equations (13.46) to (13.52): $r = 1$, $v = 6$, $\rho_{1A} = 0$, $\mu + \tau_A = 12$, $\delta_{A(B)} = 1$, $\delta_{A(C)} = 2$, $\delta_{A(D)} = 3$, $\delta_{A(E)} = 4$, $\delta_{A(F)} = 0$, $\pi_{A(BC)} = -1$, $\pi_{A(BD)} = 1$, $\pi_{A(BE)} = 0$, $\pi_{A(BF)} = 0$, $\pi_{A(CD)} = 1$, $\pi_{A(CE)} = 0$, $\pi_{A(CF)} = 0$, $\pi_{A(DE)} = -1$, $\pi_{A(DF)} = -1$, $\pi_{A(EF)} = 1$, and $\beta_{A(BCDEF)} = 2$. The responses and totals are

$Y_{\cdot A} = 12$	$Y_{\cdot A(BE)} = 22/3$	$Y_{\cdot A(CE)} = 24/3$	$Y_{\cdot A(DF)} = 15/3$
$Y_{\cdot A(BC)} = 15/3$	$Y_{\cdot A(BF)} = 14/3$	$Y_{\cdot A(CF)} = 16/3$	$Y_{\cdot A(EF)} = 23/3$
$Y_{\cdot A(BD)} = 23/3$	$Y_{\cdot A(CD)} = 25/3$	$Y_{\cdot A(DE)} = 23/3$	$Y_{\cdot A(BCDEF)} = 22/3$

$$Y_{\cdot A(B\cdot)} = 74/3 \qquad Y_{\cdot A(D\cdot)} = 86/3 \qquad Y_{\cdot A(F\cdot)} = 68/3$$

$$Y_{\cdot A(C\cdot)} = 80/3 \qquad Y_{\cdot A(E\cdot)} = 92/3 \qquad Y_{\cdot A(\cdot\cdot)} = 200/3$$

It may be verified that the above totals obtained from the observations agree with those obtained using the parameters to construct those from equations (13.49) to (13.52). Likewise, using the results from the example and solving for the parameters from (13.53) to (13.57), the original values of the parameters are obtained. Using a small example with known values is valuable for checking one's algebra and solutions, and for realistic response equations. The MAPLE program in Appendix 13.3 is useful in obtaining formulae for the effects in equations (13.54)–(13.57).

Response Equations on a Four-Plant Basis

If it is desired to analyze the data from the three types of mixtures jointly in one analysis, the yields would need to be put all on the same number of plants basis as was done in Example 13.5, that is, e.g., $Y'_{hA/3}$, $Y'_{hA(jg)}$, and $2Y'_{hA(BCDEF)}$ would all be on a four-plant basis. Y_{hA} is the yield from 12 plants and $Y_{hA(BCDEF)}$ is the yield from 2 plants. On this four-plant basis, one could reparameterize the response equations as follows:

Sole crop A

$$Y'_{hA/3} = Y_{hA} = \mu + \tau_A + \rho_{hA} + \epsilon_{hA}. \tag{13.58}$$

Mixture of three crops, A with j and g

$$Y'_{hA(jg)} = Y_{hA(jg)} = \mu + \tau_A + \rho_{hA} + \frac{1}{2}\left(\delta_{A(j)} + \delta_{A(g)}\right) + \pi_{A(jg)} + \epsilon_{hA(jg)}. \tag{13.59}$$

Mixture of all six crops, yield for crop A

$$\begin{aligned} 2Y'_{hA(BCDEF)} &= Y_{hA(BCDEF)} = \mu + \tau_A + \rho_{hA} \\ &\quad + \frac{2}{(v-1)}\sum \delta_{A(j)} + 2\beta_{A(BCDEF)} + \epsilon_{hA(BCDEF)} \\ &= \mu + \tau_A + \rho_{hA} + 2\bar{\delta}_{A(\cdot)} + 2\beta_{A(BCDEF)} \\ &\quad + \epsilon_{hA(BCDEF)}. \end{aligned} \tag{13.60}$$

Using the parameterization for the response equations can be rationalized as follows. If there were no effects from the mixture, the expected value of Y_{hA}, $Y_{hA(jg)}$, and $Y_{hA(BCDEF)}$ should be $\mu + \tau_A + \rho_{hA}$, since all responses are for four plants. Likewise, one would say, with less credibility, that ϵ_{hA}, $\epsilon_{hA(jg)}$, and $\epsilon_{hA(BCDEF)}$, as defined directly above, all have mean zero and common variance $\hat{\sigma}^2_{\epsilon A}$. The last statement can only be approximately correct since $\epsilon_{hA}/3$ from the 12-plant response equation is equal to the ϵ_{hA} from the 4-plant response equations above. Thus, one would suspect that ϵ_{hA} as defined above would have a smaller variance than the $\epsilon_{hA(jg)}$ and that $\epsilon_{hA(BCDEF)}$ would have a larger variance. If the component of variance due to variation among plants within an experimental unit is small relative to the component of variance among experimental units treated alike, then the inequality of variances will be small and, hence, can be ignored. This is what was assumed for the analyses given in Example 13.5. Thus, we shall use equations (13.59) to (13.60) for analyses of the data.

The coefficient of 1/2 in (13.59) was used because half of the four-plant area went to estimating $\delta_{A(j)}$ and the other half to estimating $\delta_{A(g)}$. Using the 1/2 puts the $\delta_{A(j)}$ and $\delta_{A(g)}$ on the same basis as the μ, τ_A, and ρ_{hA}, as well as the other parameters. In the six-crop combination, only 1/5 of the responses for crop A are used to estimate $\delta_{A(j)}$ and all $\delta_{A(j)}$ are present. Therefore, the coefficient of $\Sigma\delta_{A(j)}$ is $1/(v-1) = 1/5$ for a two-plant and 2/5 for a four-plant total, resulting in the term $2\bar{\delta}_{A(\cdot)}$.

The resulting normal equations using the model constraints

$$0 = \sum_{h=1}^{v} \rho_{hi} = 0 = \sum_{\substack{j=1 \\ \neq i,g}}^{v} \pi_{i(jg)} = \sum_{\substack{g=1 \\ \neq i,j}}^{v} \pi_{i(jg)} \text{ and } \sum_{\substack{j=1 \\ \neq i}}^{v} \delta_{i(j)} = \bar{\delta}_{i(\cdot)}(v-1)$$

are

$$Y_{\cdot i} = r(\mu + \tau_i), \tag{13.61}$$

$$Y_{\cdot i(\cdot\cdot)} = \frac{r(v-1)(v-2)}{2}\left(\mu + \tau_i + \bar{\delta}_{i(\cdot)}\right), \tag{13.62}$$

$$\begin{aligned} Y_{\cdot i(j\cdot)} &= r(v-2)(\mu+\tau_i) + \frac{r(v-1)}{2}\,\bar{\delta}_{i(\cdot)} \\ &\quad + \frac{r(v-3)}{2}\delta_{i(j)}, \end{aligned} \tag{13.63}$$

$$Y_{\cdot i(jg)} = r\left[\mu + \tau_i + \frac{1}{2}\left(\delta_{i(j)} + \delta_{i(g)}\right) + \pi_{i(jg)}\right], \tag{13.64}$$

$$Y_{\cdot i(\text{all but } i)} = r\left(\mu + \tau_i + 2\delta_{i(\cdot)} + 2\beta_{i(\text{all but } i)}\right). \tag{13.65}$$

Solutions for effects in the above normal equations are for crop $i \neq j$, g and $v = 6$:

$$\hat{\mu} + \hat{\tau}_i = \frac{Y_{\cdot i\cdot}}{r} = \bar{y}_{\cdot i\cdot}, \tag{13.66}$$

$$\hat{\bar{\delta}}_{i(\cdot)} = \frac{2Y_{\cdot i(\cdot\cdot)}}{r(v-1)(v-2)} - \bar{y}_{\cdot i} = \bar{y}_{\cdot i(\cdot\cdot)} - \bar{y}_{\cdot i}, \tag{13.67}$$

$$\begin{aligned} \hat{\delta}_{i(j)} &= \frac{2}{r(v-3)}\left[Y_{\cdot i(j\cdot)} - r(v-2)\bar{y}_{\cdot i} - \frac{r}{2}(v-1)\hat{\bar{\delta}}_{i(\cdot)}\right] \\ &= \frac{2}{3r}Y_{\cdot i(j\cdot)} - \frac{5}{3}\bar{y}_{\cdot i(\cdot\cdot)} - \bar{y}_{\cdot i}, \end{aligned} \tag{13.68}$$

$$\hat{\pi}_{i(jg)} = \frac{1}{r}\left[Y_{\cdot i(jg)} - r\bar{y}_{i\cdot} - \frac{r}{2}\left(\hat{\delta}_{i(j)} + \hat{\delta}_{i(g)}\right)\right]$$

$$= \bar{y}_{\cdot i(jg)} - \frac{1}{3r}\left[Y_{\cdot i(j\cdot)} + Y_{\cdot i(\cdot g)}\right] + \frac{5}{3}\bar{y}_{\cdot i(\cdot\cdot)}, \tag{13.69}$$

$$2\hat{\beta}_{i(\text{all but } i)} = \bar{y}_{\cdot i(\text{all the rest})} - 2\bar{y}_{\cdot i(\cdot\cdot)} + \bar{y}_{\cdot i}. \tag{13.70}$$

A numerical example for equations (13.58) to (13.70) is given below. Let the parameter values in (13.58) to (13.60) be those for the example following equation (13.57). Then, the one-replicate totals are

$$\begin{array}{llll}
Y_{\cdot A} = 12 & Y_{\cdot A(BE)} = 14.5 & Y_{\cdot A(CE)} = 15 & Y_{\cdot A(DF)} = 12.5 \\
Y_{\cdot A(BC)} = 12.5 & Y_{\cdot A(BF)} = 12.5 & Y_{\cdot A(CF)} = 13 & Y_{\cdot A(EF)} = 15 \\
Y_{\cdot A(BD)} = 15 & Y_{\cdot A(CD)} = 15.5 & Y_{\cdot A(DE)} = 14.5 & Y_{\cdot A(BCDEF)} = 20
\end{array}$$

$$\begin{array}{lll}
Y_{\cdot A(B\cdot)} = 54.5 & Y_{\cdot A(D\cdot)} = 57.5 & Y_{\cdot A(F\cdot)} = 53 \\
Y_{\cdot A(C\cdot)} = 56 & Y_{\cdot A(E\cdot)} = 59 & Y_{\cdot A(\cdot\cdot)} = 140
\end{array}$$

Again, the above totals agree with those in (13.61) to (13.65) where values of the parameters are used to obtain the totals instead of the above responses. Also, using (13.66) to (13.70) with the above totals results in solutions which are equal to the original parameters.

Appendix 13.2

Variances for Estimation of Effects Given by Equations (13.53)–(13.57)

The variances for the estimated effects given in equations (13.53) to (13.57) and Example 13.5 are

$$V\left(\hat{\mu} + \hat{\tau}_i = \bar{y}_{\cdot i}\right) = \frac{\sigma^2_{\epsilon i}}{r}, \tag{13.71}$$

$$V\left(\hat{\delta}_{i(\cdot)} = \frac{3}{4}\bar{y}_{\cdot i(\cdot\cdot)} - \frac{1}{4}\bar{y}_{\cdot i}\right) = \frac{9 \quad \sigma^2_{\epsilon i3}}{8(v-1)(v-2)} + \frac{\sigma^2_{\epsilon i}}{16r}, \tag{13.72}$$

$$\begin{aligned} V\left(\hat{\delta}_{i(j)} = \frac{3(v-2)}{2(v-3)}\bar{y}_{\cdot i(j\cdot)} - \frac{3}{4}\left(\frac{v-1}{v-3}\right)\bar{y}_{\cdot i(\cdot\cdot)} - \bar{y}_{\cdot i/4}\right) \\ = \frac{9}{8}\left(\frac{2v-5}{(v-2)(v-3)}\right)\sigma^2_{\epsilon i3} + \frac{\sigma^2_{\epsilon i}}{16r}, \end{aligned} \tag{13.73}$$

$$\begin{aligned} V\left(\hat{\pi}_{i(jg)} = \bar{y}_{\cdot i(jg)} - \left(\frac{v-2}{v-3}\right)\left(\bar{y}_{\cdot i(j\cdot)} + \bar{y}_{\cdot i(\cdot g)}\right)\right. \\ \left. + \left(\frac{v-1}{v-3}\right)\bar{y}_{\cdot i(\cdot\cdot)}\right) = \frac{(v-4)}{r(v-2)}\sigma^2_{\epsilon i3}, \end{aligned} \tag{13.74}$$

$$\begin{aligned} V\left(\hat{\beta}_{i(\text{all but } i)} = \bar{y}_{\cdot i(\text{all but } i)} - \frac{(v-1)}{4}\bar{y}_{\cdot i(\cdot\cdot)} + \frac{(v-3)}{12}\bar{y}_{\cdot i}\right) \\ = \frac{\sigma^2_{\epsilon i6}}{r} + \frac{(v-1)}{8r(v-2)}\sigma^2_{\epsilon i3} + \left(\frac{v-3}{12}\right)^2\frac{\sigma^2_{\epsilon i}}{r}, \end{aligned} \tag{13.75}$$

where $\sigma^2_{\epsilon i}$ is the error variance for sole crop i, $\sigma^2_{\epsilon i3}$ is the error variance for mixtures of three for sole crop i, and $\sigma^2_{\epsilon i6}$ is the error variance for mixtures of size six for crop i.

Variances for Estimated Effects for Equations (13.66)–(13.70)

Assuming random block effects, the variance for $\hat{\mu} + \hat{\tau}_i$ is known to be

$$V(\hat{\mu} + \hat{\tau}_i) = (\sigma^2_{\epsilon i} + \sigma^2_{\rho i}/r), \tag{13.76}$$

where $\sigma^2_{\rho i}$ is the variance component for blocks. The other variances for estimated effects given by equations (13.67) to (13.69) are

$$V\left(\hat{\delta}_{i(\cdot)} = \bar{y}_{\cdot i(\cdot\cdot)} - \bar{y}_{\cdot i}\right) = \sigma^2_{\epsilon i}\left(\frac{2}{r(v-1)(v-2)} + \frac{1}{r}\right) = \frac{11\sigma^2_{\epsilon i}}{10r}; \tag{13.77}$$

for $v = 6$ and for $\bar{y}_{\cdot i(\cdot\cdot)}$ and $\bar{y}_{\cdot i}$ independent,

$$V\left(\hat{\delta}_{i(j)} = \frac{2}{3r} Y_{\cdot i(j\cdot)} - \frac{5}{3}\bar{y}_{\cdot i(\cdot\cdot)} - \bar{y}_{\cdot i} = Y_{\cdot i(j\cdot)} \left[\frac{2}{3r} - \frac{1}{6r} = \frac{1}{2r}\right]\right.$$

$$\left. - \frac{1}{6r} \left[\frac{\text{6 other } Y_{\cdot i(jg)}}{\text{not in } Y_{\cdot i(jg)}}\right] - \bar{y}_{\cdot i}\right)$$

$$= \sigma_{\epsilon i}^2 \left(\frac{1}{r} + \frac{1}{6r} + \frac{1}{r}\right) = \frac{13\sigma_{\epsilon i}^2}{6r}; \tag{13.78}$$

$$V\left(\hat{\pi}_{i(jg)} = \frac{1}{r} Y_{\cdot i(jg)} - \frac{1}{3r}\left(Y_{\cdot i(j\cdot)} + Y_{\cdot i(g\cdot)}\right) + \frac{1}{6r} Y_{\cdot i(\cdot\cdot)}\right)$$

$$= Y_{\cdot i(jg)} \left(\frac{1}{r} - \frac{2}{3r} + \frac{1}{6r} = \frac{1}{2r}\right)$$

$$- \left(\text{6 terms } Y_{\cdot i(j'g')} \text{for } j' \neq j,\, g' \neq g\right) \left(-\frac{1}{3r} + \frac{1}{6r} = -\frac{1}{6r}\right)$$

$$+ \left(\text{3 terms } Y_{\cdot i(j^+g^+)} \text{ not considered before }\right) \left(\frac{1}{6r}\right)$$

$$= \sigma_{\epsilon i}^2 \left(\frac{1}{4r} + \frac{1}{6r} + \frac{1}{12r} = \frac{1}{2r}\right) = \frac{\sigma_{\epsilon i}^2}{2r}; \tag{13.79}$$

$$V\left(\hat{\beta}_{i(\text{all but } i)} = \frac{1}{2}\left[\frac{Y_{\cdot i(\text{all but } i)}}{r} - 2\left(\frac{Y_{\cdot i(\cdot\cdot)}}{10r}\right) + \frac{Y_{\cdot i}}{r}\right]\right)$$

$$= \sigma_{\epsilon i}^2 \left(\frac{1}{r} + \frac{4}{10r} + \frac{1}{r} = \frac{12}{5r}\right) = \frac{12\sigma_{\epsilon i}^2}{5r}. \tag{13.80}$$

As an aid in developing the above formulas, a MATHEMATICA program is given in Appendix 13.3.

Appendix 13.3

MAPLE Program as an Aid in Obtaining the Solutions Given in Equations (13.53)–(13.57)

The MAPLE program used was

```
eqs1 := {y1 = a/3, y12 = a/3 + 2*(v - 2)*j/(3*(v - 3))
y13 = a/3 + 2*(v - 2)*g/(3*(v - 3)),
y123 = a/3 + 2*(j + g)/3 + p};
a1 := solve(eqs1, {a, j, g, p});
h1 := collect(a1, {y1, y12, y13, y123}, factor);
```

MATHEMATICA Program as an Aid in Obtaining the Variances in Equations (13.78) and (13.79)

The following program using values for $v = 5, 6, 7$, and 8 was run to verify the variance formulas given by equations (13.78) and (13.79). Note that $7Er/6 + 1 = 13Er/6$ which is the variance given in (13.78) for $v = 6$. Output is also given following the program. Using a semicolon at the end of a statement gives no output, but omitting the semicolon results in output (the reverse is true for MAPLE).

```
In[123]:=
  v = 6;
  s1 = Sum[e[1, j, g], {j, 2, v - 1}, {g, j + 1, v}];
  s12 = Sum[e[1, 2, g], {g, 3, v}];
  s13 = e[1, 2, 3] + Sum[e[1, 3, g], {g, 4, v}];
  res = {e[i, j, g] e[i, j, g]- > Er, e[i, j, g] e[i, j, h]- > 0,
  e[i, j, g] e[i, f, h]- > 0, e[i, j, g] e[d, f, h]- > 0,
  e[i, j, g] e[i, f, g]- > 0};
  X1 = Simplify[2*s12/3 - 2*s1/(3*(v - 2))]
  Expand[X1^2]/.res
  X2 = Simplify[2*s13/3 - 2*s1/(3*(v - 2))];
  Expand[X2^2]/.res
  X3 = Simplify[e[1, 2, 3] - s12/3 - s13/3 + 2*s1/(3*(v - 2))]
  Expand[X3^2]/.res
```

Out[128] =

$$(3e[1,2,3] + 3e[1,2,4] + 3e[1,2,5] + 3e[1,2,6] - e[1,3,4] - e[1,3,5] - e[1,3,6] - e[1,4,5] - e[1,4,6] - e[1,5,6])/6$$

Out[129] =
$$\frac{7\mathrm{Er}}{6}$$

Out[131] =
$$\frac{7\mathrm{Er}}{6}$$

Out[132] =
$$(3e[1,2,3]-e[1,2,4]-e[1,2,5]-e[1,2,6]-e[1,3,4]-e[1,3,5]$$
$$-e[1,3,6]+e[1,4,5]+e[1,4,6]+e[1,5,6])/6$$

Out[133] =
$$\frac{\mathrm{Er}}{2}$$

CHAPTER 14

Varying Densities for Some or All Crops in a Mixture

14.1 Introduction

The simplest form of intercropping with three or more crops in the mixture and with one major crop and two or more minor crops was considered in Chapter 12. The complexity of the statistical analyses over that in Chapter 2 (two crops) of Volume I is increased. The methods of Chapters 3 and 4 of Volume I were extended to mixtures of three or more crops in Chapter 13. Analyses for individual crop responses for each crop as well as analyses for combined responses for all crops in the mixture are presented. The density for a given crop in the mixture was held constant from mixture to mixture. In the present chapter, cropping systems which allow varying densities for some or all crops are considered. The methods presented herein are a generalization of those presented in Chapter 5 of Volume I.

Many patterns for varying and/or constant densities in a mixture are possible. The particular densities selected for study will depend on the makeup of the crop mixture as well as the goals of the experiment. With one major crop and two or more minor crops:

(i) The density of the major crop could be varied and the densities of the minor crops kept constant.
(ii) The density of the major crop could be held constant and some or all of the densities of the minor crops varied.
(iii) The densities of all crops in the mixture could be varied.

With three or more major crops and with some or no minor crops in a mixture, the following situations are possible:

(i) The densities of all major crops in the mixture could be varied.
(ii) The densities of two or more major crops could be constant and the densities of the remaining crops could be varied.
(iii) The densities of any minor crops included in (i) or (ii) could be varied or held constant.

As discussed in Chapter 5 of Volume I, serious attention needs to be given to selecting the various density levels for each crop. The experimenter needs to decide whether to make the levels selected for one crop dependent or independent of the levels selected for the remaining crops in a mixture. It may make sense to approach a maximum density for all crops in the mixture as the total number of plants, regardless of crop, is the total population level beyond which there will be no increased yields. The amount of moisture, plant nutrients, sunlight, etc. may dictate the maximum population level that can be supported on a plot of ground. It is well known that overpopulation can result in reduced or even zero yields. In order to pinpoint density levels producing maximum or near-maximum responses, it is advisable to select levels somewhat beyond the level giving maximum response. For example, the maximum yield of maize may be attained with 60,000 plants per hectare. A level of 70,000, or even 80,000, plants per hectare should result in decreased yields and should be included for study in an experiment. In determining response curves, experimenters often make the mistake of including only levels which "would be used in practice." The inclusion of levels beyond those normally used in practice results in a more accurate response curve showing the relationship between response and density level. If a response curve does not show a decrease at the highest density, it is not clear that the maximum has been attained and that higher density levels should have been included.

14.2 Treatment Design

Several treatment designs may be used for studying responses over varying density levels of the crops in the mixture. We shall list some of the possible designs for studying yield–density relationships.

14.2.1 Design 1

Consider a mixture of three crops at densities $0 < d_{i1} < d_{i2} < \cdots < d_{in_i}$ for crop i at n_i density levels. Then, for $n_1 = 3$, $n_2 = 2$, and $n_3 = 4$, the following combinations (marked X) are obtained where, for example, crop one is cassava, crop two is beans, and crop three is maize:

Crop one (cassava)

Crop three (maize)	d_{11} Crop two (beans)		d_{12} Crop two (beans)		d_{13} Crop two (beans)	
	d_{21}	d_{22}	d_{21}	d_{22}	d_{21}	d_{22}
d_{31}	X	X	X	X	X	X
d_{32}	X	X	X	X	X	X
d_{33}	X	X	X	X	X	X
d_{34}	X	X	X	X	X	X

In addition to the above 24 combinations, the 3 crops as sole crops could be included to obtain $(3\times2\times4)+(3+2+4) = 33$ entries. Here the lowest densities $d_{i1}, i = 1, 2,$ and 3, are greater than zero. As the number of density levels for a crop and the number of crops increase, the total number of entries for an experiment increases rapidly. For example, including a fourth crop at three density levels, say, to the above set would result in $(3 \times 2 \times 4 \times 3) + (3 + 2 + 4 + 3) = 84$ entries. Therefore, the experimenter needs to exercise considerable care in selecting the precise levels and their number in order that the number of entries does not go beyond what can be done experimentally. This treatment design contains all possible combinations of density levels plus the levels for each of the sole crops.

14.2.2 *Design 2*

A procedure for reducing the number of entries and the space requirements would be to utilize the ideas of Federer and Scully (1993) in the manner shown in Figure 14.1, using the previous example for three crops. There are n_1 experimental units in each replicate; the density for crop two varies from the lowest to the highest density horizontally either continuously increasing or increasing by increments; and the crop three densities vary in the same way but vertically, with highest combined densities being in the lower right-hand corner of an experimental unit. The experimenter would divide each experimental unit into n_2n_3 equal-sized rectangles and obtain the response for each of these rectangles. The density level for each rectangle would be the average density in that rectangle. The n_1 experimental units are randomly allocated in each replicate in the experiment and the crop two and

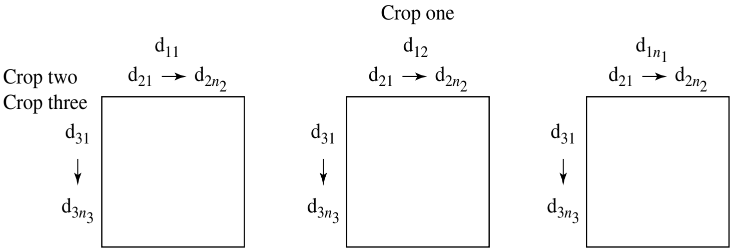

FIGURE 14.1. Schematic plan for Design 2 for one replicate.

crop three densities are systematically increasing within the experimental unit. Thus, the crop one density levels are somewhat akin to a "whole plot" and the density levels of crops two and three are somewhat like "split plots." A response function, e.g., a second-degree polynomial, would be fitted, the maximum value on the response surface, and/or the area under the response function could be used as the response for the experimental unit (see Federer and Scully, 1993). The selection of which crop to use as crop one is important, but probably one crop would be an obvious candidate. For the first example above, cassava would be crop one because a large experimental unit relative to the one needed for maize or beans would be required. In other situations, one of the crops may utilize well-defined discrete levels and, hence, would be a candidate to be crop one. There should be no gradients within each of the experimental units in order that a gradient effect does not become confounded with the effect of density level on the response.

14.2.3 Design 3

If the densities of four crops in a mixture are to be varied, one suggested procedure is to use all n_1n_2 levels for crops one and two. Then, for each combination, use the experimental unit described above for three crops. This would result in the plan given in Figure 14.2. The n_1n_2 experimental units would be randomized in the experiment and each experimental unit would be divided up into n_3n_4 equal-sized rectangles with a response and density level being obtained for each rectangle. The predicted values from the fitted response surface for each experimental unit or some other statistic would be used in an analysis of the data.

14.2.4 Design 4

Still another parsimonious treatment design for studying density-yield relationships for c crops at varying density levels is to use 2^c combinations, where 0

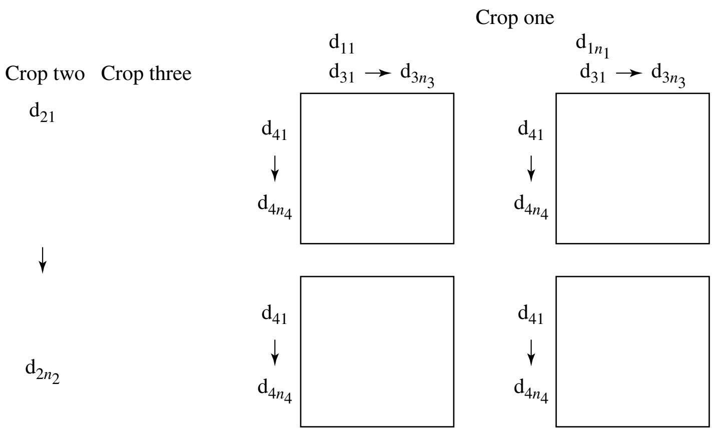

FIGURE 14.2. Schematic plan for Design 3 in one replicate.

means density constant for crop i, $i = 1, \ldots, c$, and 1 means increasing densities for the designated crop. The $000\cdots$ combination is the lowest density level for all crops and crop one, for example, is at this lowest level when crop two, say, is increasing within the experimental unit. There will be 2^c experimental units in each replicate of this type of experiment.

14.2.5 Design 5

Another procedure for reducing the number of combinations and space requirements is to utilize the ideas from fractional replication of complete factorials (see Raktoe *et al.*, 1981). The above procedure for Design 1 includes all possible combinations from the factorial. Saturated main effects fractions or Resolution V, fractions which allow for the estimation of all two-factor interactions, may satisfy the requirements of an experimenter. Certainly in the preliminary stages of studying yield–density relationships, one of these fractions usually will be satisfactory. A Resolution IV fraction allows the estimation of main effects and of sums of two-factor interactions. A Resolution VI fraction allows the estimation of all main effects, of all two-factor interactions, and sums of three-factor interactions, and a Resolution VII fraction allows estimation of all main effects, all two-factor interactions, and all three-factor interactions. Higher-ordered interaction terms are assumed to be zero. A saturated fractional replicate contains as many observations as there are effects. The construction of fractional factorial treatment designs for the general factorial is an unsolved problem. Usually, a computer search will be made to obtain a relatively good fraction which can be evaluated against the best possible one (see Anderson and Federer, 1995). Given that there will be three levels of each of c crops, the number of observations needed for designs of various resolutions are listed in Figure 14.3. As can be seen from the figure, the number of combinations increases rapidly as the resolution and the number of crops increase. In the general case of c crops with n_i levels for crop i, a saturated Resolution III fractional replicate requires $1 + \sum_{i=1}^{c}(n_i - 1)$ observations, a saturated Resolution V fractional replicate requires

$$1 + \sum_{i=1}^{c}(n_i - 1) + \sum_{i<j}\sum(n_i - 1)(n_j - 1)$$

Crops	III	IV	V	VI	VII	Full
3	7	14	19	38	27	27
4	9	18	33	66	65	81
5	11	22	51	102	131	243
6	13	26	73	146	233	729
7	15	30	99	198	379	2187

FIGURE 14.3. Number of combinations.

observations, and a saturated Resolution VII fractional replicate requires

$$1+\sum_{i=1}^{c}(n_i-1)+\sum_{i<j}\sum(n_i-1)(n_j-1)$$
$$+\sum_{i<j<k}\sum\sum(n_i-1)(n_j-1)(n_k-1)$$

observations. Many other fractions are possible such as fractions for estimating the linear, quadratic, and only linear-by-linear interactions.

Any saturated fraction can easily be obtained using the one-at-a-time procedure studied by Anderson and Federer (1975, 1995). This is the worst possible design variancewise for main effects, Resolution III plans but gets better as the Resolution increases. To illustrate this method, consider $c = 3$ crops each at $n = 3$ levels for a saturated Resolution V fractional replicate, where each row represents a combination of the three crop densities and the corresponding X^* matrix (see Anderson and Federer, 1995) are given in Figure 14.4.

The determinant of this matrix is $19,683 = 3^9$, whereas from the Anderson-Federer (1995) article we note that the determinant of X^* lies between 1 and $N^{N/2}/s^{n/2}$, where N is the size of the fraction, here 19, s is the number of levels of a factor, here 3, and n is the number of factors, here taken to be $(19-1)/(s-1) = 9$. This upper limit is not reachable because the conditions of Theorem 3.3 of Anderson and Federer (1995) cannot be satisfied for this fraction. The value of the

Combination	Resulting X^* matrix			
000	1000000	0000	0000	0000
100	1100000	1010	1010	0000
200	1010000	0101	0101	0000
010	1001000	1001	0000	1010
020	1000100	0110	0000	0101
001	1000010	0000	1001	1001
002	1000001	0000	0110	0110
110	1101000	0100	1010	1010
120	1100100	0001	1010	0101
210	1011000	0010	0101	1010
220	1010100	1000	0101	0101
101	1100010	1010	0100	1001
102	1100001	1010	0001	0110
201	1010010	0101	0010	1001
202	1010001	0101	1000	0110
011	1001010	1001	1001	0100
012	1001001	1001	0110	0001
021	1000110	0110	1001	0010
022	1000101	0110	0110	1000

FIGURE 14.4. Combinations and X^* matrix for a Resolution V fraction for three crops at three density levels each.

determinant here is considerably above the lower limit of 1. Three-factor interaction terms are confounded (aliased) with the main effects and the two-factor interaction effects. If the three-factor interaction effects are negligible or nonexistent, then the estimates of the remaining effects are unbiased.

Any number of levels of densities for the different crops may be used with the above construction procedure. Note that for two-factor interactions, all possible combinations for any two crops are included. If certain kinds of two-factor interactions can be considered to be unimportant, then the size of the fraction may be reduced by eliminating the required combinations.

14.3 Statistical Analyses for Sole Crop Response

For crop h in a mixture of c crops in Design 1 and for the v entries arranged in a randomized complete block experiment design of r complete blocks, let the response equation, for crop $h = 1$ say, be of the usual form:

$$Y_{gd_{1i}(d_{2i}d_{3i}\cdots)} = \mu + \rho_g + \tau_{d_{1i}(d_{2i}d_{3i}\cdots)} + \epsilon_{gd_{1i}(d_{2i}d_{3i})}, \tag{14.1}$$

where μ is a general mean effect, ρ_g is the gth complete block effect, $\tau_{d_{1i}(d_{2i}d_{3i}\cdots)}$ is the effect of the $d_{1i}(d_{2i}d_{3i}\cdots)$th combination from the $n_1 \times n_2 \times n_3 \times \cdots$ combinations of the density levels $d_{1i}, d_{2i}, d_{3i}, \ldots, i = 1, 2, \ldots, n_h, h = 1, 2, \ldots, c$, the number of crops in a mixture, and the $\epsilon_{gd_{1i}(d_{2i}d_{3i}\cdots)}$ are random error terms distributed with zero mean and variance $\sigma^2_{\epsilon h}$. An analysis of variance (ANOVA) for this situation is given in Table 14.1 for the case when the lowest density level is not zero. The sums of squares are computed in the usual manner for a factorial treatment design in a randomized complete block experiment design.

If desired, each of the treatment sums of squares could be partitioned into single degree of freedom contrasts such as linear, quadratic, etc., or some other set of

TABLE 14.1. ANOVA for Responses Using Response Model (14.1)

Source of variation	d.f	SS	MS
Total	$rn_1n_2n_3\cdots = rv$		
Correction for mean	1		
Complete blocks = R	$r - 1$		
Levels of crop one = C1	$n_1 - 1$		
Levels of crop two = C2	$n_2 - 1$		
Levels of crop three = C3	$n_3 - 1$		
C1 × C2	$(n_1 - 1)(n_2 - 1)$		
C1 × C3	$(n_1 - 1)(n_3 - 1)$		
C2 × C3	$(n_2 - 1)(n_3 - 1)$		
C1 × C2 × C3	$(n_1 - 1)(n_2 - 1)(n_3 - 1)$		
Residual error	$(r - 1)(n_1n_2n_3\cdots - 1)$ $= (r - 1)(v - 1)$		

contrasts dependent on the particular response function used for the relationship between a response such as yield and density level.

A response function of the following nature might be suitable for the $n_2 n_3$ responses obtained on a single experimental unit gd_{1h} from Design 2:

$$Y_{ij} = \alpha + \beta_1 d_{2i} + \beta_2 d_{2i}^2 + \beta_3 d_{3j} + \beta_4 d_{3j}^2 + \beta_5 d_{2i} d_{3j} + \beta_6 d_{2i} d_{3j}^2 + \beta_7 d_{2i}^2 d_{3j} + \epsilon_{ij}, \quad (14.2)$$

where $i = 1, \ldots, n_2$, $j = 1, \ldots, n_3$, $h = 1, \ldots, n_1$, the β's are polynomial regression coefficients, and the d_{2i} and d_{3j} are the various density levels for crops two and three. Of course, other response functions may be more appropriate than the above one. However, this particular model does allow for linear and curvilinear responses for each crop as well as for some rather well-behaved interaction terms. The responses used would be the predicted values from the above regression equation. This response model equation may also be used for each of the $n_1 n_2$ experimental units from Design 3. The main object of this analysis is to show the effect of changing levels of the densities of crops two and three at each level of crop one. An alternate analysis would be a MANOVA (multivariate analysis of variance) or discriminant function analysis, using the seven regression coefficients as the seven variates and determining their effects over all levels of crop one. Still another analysis would be to obtain the estimated maximum responses from the regression function in equation (14.2) in each of the rn_1 experimental units and perform an analysis on these values.

The analysis for Design 3 follows that given in Table 14.2 except that there are $n_1 n_2$ experimental units and levels in each replicate instead of n_1. In addition, the seven regression coefficients may be computed for levels of crop 1 summed over levels of crop 2, for levels of crop 2 summed over levels of crop 1, and for the crop 1 by crop 2 interaction with $n_1 n_2 - n_1 - n_2 + 1$ degrees of freedom (by subtraction).

The analysis for Design 4 in an RCBD (randomized complete block design) follows that for a standard 2^c factorial treatment design, and an ANOVA is presented in Table 14.3 for $c = 3$ crops. The effect measured here is whether or not density is increasing in an experimental unit. Note that there are experimental units with constant density for all crops of the mixture, experimental units with one of the c crops having increasing density and the other crops at a constant density, experimental units with two crops having increasing densities and the rest at a constant density, etc. When using the total yield from an experimental unit, the effect being observed is that of increasing density for one or more crops. When the experimental unit is subdivided and responses obtained, other statistics that may hold interest for the experimenter are as follows:

(i) the maximum response,
(ii) the area under a response curve,
(iii) the linear regression coefficient, and/or
(iv) the linear and quadratic regression coefficients.

TABLE 14.2. ANOVA for r Replicates of Design 2 in an RCBD for Equation (14.2).

Source of variation	d.f.	SS	MS
Crop one levels			
Replicates = R	$r - 1$		
Crop one levels = C1	$n_1 - 1$		
C1 $\times R$	$(n_1 - 1)(r - 1)$		
Crops two and three			
Regression coefficients	$7rn_1$		
Regression coefficients at level d_{11}	7		
Crop two linear, β_1	1		
Crop two quadratic, β_1	1		
Crop three linear, β_3	1		
Crop three quadratic, β_4	1		
Linear by linear, β_5	1		
Linear by quadratic, β_6	1		
Quadratic by linear, β_7	1		
Regression coefficients at level d_{12}	7		
Crop two linear, β_1	1		
Crop two quadratic, β_2	1		
Crop three linear, β_3	1		
Crop three quadratic, β_4	1		
Linear by linear, β_5	1		
Linear by quadratic, β_6	1		
Quadratic by linear, β_7	1		
...			
Regression coefficients at level d_{1n_1}	7		
Crop two linear, β_1	1		
Crop two quadratic, β_2	1		
Crop three linear, β_3	1		
Crop three quadratic, β_4	1		
Linear by linear, β_5	1		
Linear by quadratic, β_6	1		
Quadratic by linear, β_7	1		
Regression coefficients for all	1		
levels of $d_{1i} \times$ replicates	$7n_1(r - 1)$		

These responses would then be used in an ANOVA or other statistical analysis.

For Design 5, it is recommended that a fractional replicate not be replicated. Instead, the fraction should be completed to a full factorial before any replication is performed. The reason for this is that the resulting orthogonality and information on additional effects will make this a much more efficient procedure. Information on the size of the effects assumed to be zero in the fractional replicate can be obtained from the complete factorial. In preliminary yield–density studies, replication is often not needed and the fractional replicate will be sufficient to provide the preliminary results for finding the range in which to perform a more compre-

TABLE 14.3. ANOVA for Treatment Design 4 in an RCBD.

Source of variation	d.f.	SS	MS
Total	$r2^c$		
Correction for mean	1		
Replicate R	$(r-1)$		
Treatment T	(2^c-1)		
Crop one	1		
Crop two	1		
Crop one × Crop two	1		
Crop three	1		
Crop one × Crop three	1		
Crop two × Crop three	1		
Crop one × Crop two × Crop three	1		
etc.			
$R \times T$	$(r-1)(2^c-1)$		

hensive and properly replicated experiment over time and space. For unreplicated fractional or complete factorials, use may be made of Daniel's (1959) half-normal plot procedure to perform an analysis and obtain confidence intervals. He considers only two levels, but the procedure works for any number of levels (see Birnbaum, 1959, Krane, 1963). Simply compute the sums of squares for all the single degree of freedom contrasts and use the square roots of these sums of squares in the manner described by Daniel (1959). An ANOVA for the fraction given in Figure 14.4 is presented in Table 14.4. Here, we have considered there to be a single replicate of the $N = 19$ combinations for this Resolution V treatment design. Usually, the experimenter would want to compute, and should, the sums of squares for single degree of freedom contrasts. Then, using the square roots of the sums of squares for the 18 single degree of freedom contrasts, the half-normal probability plot method of Daniel (1959) may be used to construct confidence intervals on the effects even though there was no replication of the treatment combinations. This method assumes that some of the individual degree of freedom contrasts represent error contrasts.

14.4 Statistical Analyses for Combined Responses from All Crops

Instead of considering individual crop responses such as was done in the previous section, the responses from the k crops could be combined into a created variable such as

(i) total monetary value of the k crop responses,
(ii) total calories of the k crop responses,
(iii) total protein of the k crop responses,

TABLE 14.4. ANOVA for a Fractional Replicate with $N = 19$ Combinations.

Source of variation	d.f.	SS	MS
Total	19		
Correction for mean	1		
Crop one = A	2		
Crop two = B	2		
Crop one = C	2		
A × B	4		
A × C	4		
B × C	4		

(iv) land use efficiency as measured by a land equivalent ratio, and/or
(v) some other function of the k responses.

Using these created variables, ANOVAs like those given above can be obtained. Such analyses as these are more useful than those in the previous section, as they deal with the system of intercropping rather than concentrating on the components, individual crops, of the system.

Other analyses such as AMMI (additive main effects and multiplicative interaction; see Gauch, 1988, Gauch and Zobel, 1988, Ezumah *et al.*, 1991, and related references) and MANOVA (multivariate analysis of variance, see Chapter 4 of Volume I, e.g.) may be useful in certain cases. For these analyses, the responses for the individual crops form the variates for the multivariate analyses, and functionals combining response from all crops would be obtained. The interpretation of the resulting principle components and canonical variates may be a problem. These statistics may differ if a logarithmic or some other transformation of the responses had been made before using AMMI or MANOVA. Hence, selective and careful use of these procedures are necessary in order for them to be of practical and interpretive usefulness for a researcher. In some cases, little, if anything new, is added by these more complex procedures (see, e.g., Ezumah *et al.*, 1991). For many situations using analyses involving the created variables in (i), (ii), (iii), and (iv) above will suffice. In some cases, other functionals such as AMMI and MANOVA may be useful.

14.5 Modeling Responses in Sole Crop Yields

Many yield–density or response–density relationships can be formulated [see Section 3 of Morales (1993) and references therein]. In order not to make the modeling process too complicated, we shall consider simple relationships. A simple model is a linear relation between yield and density. For a randomized complete block design with three sole crops, say cassava = c, maize = m, and beans = b, the response models are

$$Y_{gch} = \mu_{\cdot c} + \rho_{gc} + \beta_{1c}(d_{ch} - \bar{d}_{c\cdot}) + \epsilon_{gch} = \beta_{0gc} + \beta_{1c}d_{ch} + \epsilon_{gch}, \tag{14.3}$$

$$Y_{gmi} = \beta_{0gm} + \beta_{1m}d_{mi} + \epsilon_{gmi}, \tag{14.4}$$

$$Y_{gbj} = \beta_{0gb} + \beta_{1b}d_{bj} + \epsilon_{gbj}. \tag{14.5}$$

where $g = 1, 2, \ldots, r$, $h = d_{c1}, d_{c2}, \ldots, d_{cn_c}$ for cassava, $i = d_{m1}, d_{m2}, \ldots, d_{mn_m}$ for maize, $j = d_{b1}, d_{b2}, \ldots, d_{bn_b}$ for beans, μ_c = common mean for cassava, ρ_{gc} = gth complete block effect, $\bar{d}_{c\cdot}$ = average density; the $\beta_{0g\cdot} = \mu_c + \rho_{gc} - \beta_{1c}\bar{d}_{c\cdot}$ are the intercepts for a crop $(\cdot)$ for the gth complete block, the $\beta_{1\cdot}$'s are the linear regression coefficients for each crop $(\cdot)$, ϵ_{gch} is a random error effect for cassava distributed with mean zero and variance $\sigma^2_{\epsilon c}$, ϵ_{gmi} is a random error effect for maize distributed with mean zero and variance $\sigma^2_{\epsilon m}$, and ϵ_{gbj} is a random error effect for beans distributed with mean zero and variance $\sigma^2_{\epsilon b}$.

As explained in Chapter 5 of Volume I, other simple or more complex yield–density models may be used in place of models (14.3)–(14.5). We shall use the above models since they are simple and illustrate the procedure. For this situation, the least squares solutions for the various parameters for three-crop mixtures of c=cassava, m=maize, and b=beans are

$$\hat{\beta}_{0gc} = \bar{y}_{gc\cdot} - \bar{d}_{c\cdot}\hat{\beta}_{1gc}, \tag{14.6}$$

$$\hat{\beta}_{0gm} = \bar{y}_{gm\cdot} - \bar{d}_{m\cdot}\hat{\beta}_{1gm}, \tag{14.7}$$

$$\hat{\beta}_{0gb} = \bar{y}_{gb\cdot} - \bar{d}_{b\cdot}\hat{\beta}_{1gb}, \tag{14.8}$$

$$\hat{\beta}_{1gc} = \sum_{h=1}^{n_c}(d_{ch} - \bar{d}_{c\cdot})(Y_{gch} - \bar{y}_{gc\cdot}) \Big/ \sum_{h=1}^{n_c}(d_{ch} - \bar{d}_{c\cdot})^2, \tag{14.9}$$

$$\hat{\beta}_{1gm} = \sum_{i=1}^{n_m}(d_{mi} - \bar{d}_{m\cdot})(Y_{gmi} - \bar{y}_{gm\cdot}) \Big/ \sum_{i=1}^{n_m}(d_{mi} - \bar{d}_{m\cdot})^2, \tag{14.10}$$

$$\hat{\beta}_{1gb} = \sum_{j=1}^{n_b}(d_{bj} - \bar{d}_{b\cdot})(Y_{gbj} - \bar{y}_{gb\cdot}) \Big/ \sum_{j=1}^{n_b}(d_{bj} - \bar{d}_{b\cdot})^2, \tag{14.11}$$

$$\hat{\beta}_{0c} = \bar{y}_{\cdot c\cdot} - \bar{d}_{c\cdot}\hat{\beta}_{1c}, \tag{14.12}$$

$$\hat{\beta}_{0m} = \bar{y}_{\cdot m\cdot} - \bar{d}_{m\cdot}\hat{\beta}_{1m}, \tag{14.13}$$

$$\hat{\beta}_{0b} = \bar{y}_{\cdot b\cdot} - \bar{d}_{b\cdot}\hat{\beta}_{1b}, \tag{14.14}$$

$$\hat{\beta}_{1c} = \sum_{h=1}^{n_c}(d_{ch} - \bar{d}_{c\cdot})(\bar{y}_{\cdot ch} - \bar{y}_{\cdot c\cdot}) \Big/ \sum_{h=1}^{n_c}(d_{ch} - \bar{d}_{c\cdot})^2, \tag{14.15}$$

$$\hat{\beta}_{1m} = \sum_{i=1}^{n_m}(d_{mi} - \bar{d}_{m\cdot})(\bar{y}_{\cdot mi} - \bar{y}_{\cdot m\cdot}) \Big/ \sum_{i=1}^{n_m}(d_{mi} - \bar{d}_{m\cdot})^2, \tag{14.16}$$

$$\hat{\beta}_{1b} = \sum_{j=1}^{n_b}(d_{bj} - \bar{d}_{b\cdot})(\bar{y}_{\cdot bj} - \bar{y}_{\cdot b\cdot}) \Big/ \sum_{j=1}^{n_b}(d_{bj} - \bar{d}_{b\cdot})^2, \tag{14.17}$$

where the $\bar{y}$ and $\bar{d}$ are mean values for the corresponding values of responses and densities, respectively, and β_{0c}, β_{0m}, and β_{0b} are the intercepts averaged over replicates. Extension of the above to c more than 3 crops is straightforward. The above assumes that density levels are the same from replicate to replicate. If this is not the true situation, e.g., missing plots occur, the above formulas will need to be adjusted to account for the change in density levels.

In the event that monocrop responses are not available, it is possible to model yield–density relationships using the lowest density levels for all other crops but the one under consideration. For example, this relationship for crop one, say, is obtained from the responses at levels $d_{21}d_{31}\cdots$ for crops two, three, etc. Using the above cassava–maize–bean example, the yield–density models would be

$$Y_{gd_{ch}(d_{m1}d_{b1})} = \beta_{g0c} + \beta_{1c}d_{ch} + \epsilon_{gd_{ch}(d_{m1}d_{b1})}, \tag{14.18}$$

$$Y_{gd_{mi}(d_{c1}d_{b1})} = \beta_{g0m} + \beta_{1m}d_{mi} + \epsilon_{gd_{mi}(d_{c1}d_{b1})}, \tag{14.19}$$

$$Y_{gd_{bj}(d_{c1}d_{m1})} = \beta_{g0b} + \beta_{1b}d_{bj} + \epsilon_{gd_{bj}(d_{c1}d_{m1})}, \tag{14.20}$$

where $gd_{ch}(d_{m1}d_{b1})$ is for level d_{ch} at levels d_{m1} and d_{b1} in replicate g and where the regression coefficients are defined in a manner similar to that for equations (14.3)–(14.5). The least squares solutions for the parameters of (14.18)–(14.20) are much the same as given in equations (14.6)–(14.17). Because of the direct application of the above solution with the necessary changes to account for computing the regressions on the lowest-density levels of all crops but the one in question, the least squares solutions are not given, as they are straightforward.

14.6 Modeling Responses for Mixtures Based on Sole Crop Model

In modeling responses for yield–density relations for mixtures of k crops, we shall follow the format of Chapter 5 of Volume I in that the model for sole crops will be extended to include a term for the effect of the mixture at the particular densities of the mixture. For the three-crop mixture example used for equations (14.3), (14.4), and (14.5), the three model equations are

$$\begin{aligned} Y_{gc(m,b)d_{ch}d_{mi}d_{bj}} &= \beta_{0gc} + \beta_{1c}d_{ch} \\ &\quad + \gamma_{c(m,b)}(d_{ch}d_{mi}d_{bj}) + \epsilon_{gc(m,b)d_{ch}d_{mi}d_{bj}}, \end{aligned} \tag{14.21}$$

$$\begin{aligned} Y_{gm(c,b)d_{ch}d_{mi}d_{bj}} &= \beta_{0gm} + \beta_{1m}d_{mi} \\ &\quad + \gamma_{m(c,b)}(d_{ch}d_{mi}d_{bj}) + \epsilon_{gm(c,b)d_{ch}d_{mi}d_{bj}}, \end{aligned} \tag{14.22}$$

$$Y_{gb(c,m)d_{ch}d_{mi}d_{bj}} = \beta_{0gb} + \beta_{1b}d_{bj}$$

$$+ \gamma_{b(c,m)}(d_{ch}d_{mi}d_{bj}) + \epsilon_{gb(c,m)d_{ch}d_{mi}d_{bj}} \tag{14.23}$$

Let $\gamma_{c(m,b)}(d_{ch}d_{mi}d_{bj}) = \gamma_{c(m,b)}(hij)$ for simplicity, and similarly for the $\gamma_{m(c,b)}(hij)$ and $\gamma_{b(c,m)}(hij)$ terms. The other symbols are defined above.

Solutions for the $\gamma_{c(mb)}(hij)$, $\gamma_{m(cb)}(hij)$, and $\gamma_{b(cm)}(hij)$ parameters are

$$\hat{\gamma}_{c(mb)}(hij) = \bar{y}_{\cdot c(mb)}(hij) - (d_{ch} - \bar{d}_{c\cdot})\hat{\beta}_{1c} - \bar{y}_{\cdot c\cdot}, \tag{14.24}$$

$$\hat{\gamma}_{m(cb)}(hij) = \bar{y}_{\cdot m(cb)}(hij) - (d_{mi} - \bar{d}_{m\cdot})\hat{\beta}_{1m} - \bar{y}_{\cdot m\cdot}, \tag{14.25}$$

$$\hat{\gamma}_{b(cm)}(hij) = \bar{y}_{\cdot b(cm)}(hij) - (d_{bj} - \bar{d}_{b\cdot})\hat{\beta}_{1b} - \bar{y}_{\cdot b\cdot}, \tag{14.26}$$

where $\bar{y}_{\cdot c(mb)}(hij)$ is the mean value over replicates of responses for cassava at density level $h = d_{ch}$ in the presence of maize at density level $i = d_{mi}$ and beans at density level $j = d_{bj}$. The other symbols are as described previously.

In order for the solutions in (14.6) to (14.17) to be least squares solutions, it is necessary for the error variance of the ϵ_{gch} values, say, to be identically and independently distributed for each density level $h = d_{ch}$ and for each value of g. If the error variance varies with density level, then a weighted least squares procedure should be used. Likewise, in computing the variances for the $\hat{\gamma}$ values, the variance for a crop in monoculture and that for the same crop in an intercrop mixture will be assumed to be the same. If this assumption is not tenable, then the $\hat{\gamma}_{c(mb)}(hij)$ in (14.24) will have one variance for $\bar{y}_{\cdot c(mb)}(hij)$ and a different variance for $\hat{\beta}_{1c}$ and $\bar{y}_{\cdot c\cdot}$, which brings in the Behrens–Fisher situation of unequal variances. However, if the degrees of freedom for the error variances are relatively large, Grimes and Federer (1984) have demonstrated that the equal-variance situation may be used without losing much power. If $\sigma^2_{\epsilon c}$ is the error variance for cassava in monoculture and $\sigma^2_{\epsilon c(mb)}$ is the error variance for cassava in polyculture, the intercrop mixture, then the error variance for $\hat{\gamma}_{c(ma)}(hij)$ is

$$\begin{aligned} V\left[\hat{\gamma}_{c(mb)}(hij)\right] &= \frac{\sigma^2_{\epsilon c(mb)}}{r} + \frac{\sigma^2_{\epsilon c}}{r}\left[\frac{1}{n_c} + \frac{(d_{ch} - \bar{d}_{c\cdot})^2}{\sum_{h=1}^{n_c}(d_{ch} - \bar{d}_{c\cdot})^2}\right] \\ &= \frac{\sigma^2_{\epsilon c}}{r}\left(1 + \frac{1}{n_c} + \frac{(d_{ch} - \bar{d}_{c\cdot})^2}{\sum_{h=1}^{n_c}(d_{ch} - \bar{d}_{c\cdot})^2}\right) \end{aligned} \tag{14.27}$$

when $\sigma^2_{\epsilon c} = \sigma^2_{\epsilon c(mb)}$. Similarly, the error variances for $\hat{\gamma}_{m(cb)}(hij)$ and $\hat{\gamma}_{b(cm)}(hij)$ values are

$$V\left[\hat{\gamma}_{m(cb)}(hij)\right] = \frac{\sigma^2_{\epsilon m}}{r}\left[1 + \frac{1}{n_m} + \frac{(d_{mi} - \bar{d}_{m\cdot})^2}{\sum_{i=1}^{n_m}(d_{mi} - \bar{d}_{m\cdot})^2}\right] \tag{14.28}$$

and

$$V\left[\hat{\gamma}_{b(cm)}(hij)\right] = \frac{\sigma^2_{\epsilon b}}{r}\left[1 + \frac{1}{n_b} + \frac{(d_{bj} - \bar{d}_{b\cdot})^2}{\sum_{j=1}^{n_b}(d_{bj} - \bar{d}_{b\cdot})^2}\right]. \tag{14.29}$$

If error variances for monoculture and polyculture differ, then (14.28) and (14.29) should be put in the form of the middle term of (14.27). When equations (14.18)–

(14.20) are used, it is more unlikely that heterogeneous error variances will be encountered.

The above error variances may be used to test the hypothesis that $\gamma_{c(mb)}(hij)$, $\gamma_{m(ch)}(hij)$, or $\hat{\gamma}_{b(cm)}(hij)$ is equal to zero. Also, linear contrasts among the estimates of the above parameters may be tested. The $\hat{\gamma}_{c(mb)}(hij)$ values may be partitioned in the manner of a three-factor factorial if desired. Patterns among these estimates may also be of interest to the researcher.

An analysis of variance to utilize in testing the null hypotheses about the parameter value in the various models may be performed as shown in Table 5.3 of Volume I. The application is straightforward and is not repeated here. The "Error from regression" and "Biblends" (here "Multiblends") sums of squares may be partitioned into sums of squares for more complex models and for various patterns among the γ values. In addition, the results from the Appendix to Chapter 5 of Volume I may be used to perform additional tests for significance for intercropped mixtures of three or more crops. Most procedures are straightforward and, hence, are not repeated here.

The above discussion has concentrated on treatment design 1 from Section 14.3. The application to Designs 2, 4, and 5 is straightforward, but note that a weighted regression approach will need to be used for Designs 2 and 3. The regressions for crops two and three in the three-crop mixture have a smaller variance, i.e.; there are n_1r replicates for the regression coefficients for densities of crops two and three and crop one has r replicates of these coefficients. Similarly for Design 3, the regression coefficients from crops three and four have rn_1n_2 estimates, whereas crops one and two have rn_2 or rn_1 estimates, respectively.

The following example is presented to expand and clarify the ideas presented here.

14.7 An Example

Data for maize and bean yields at the zero density of melon were obtained from Aidar (1978). In order to obtain an example for mixtures of three crops with varying densities of all three crops (an example was not found in the literature; see Chapter 20), it was necessary to construct one. The data in the last six columns of Table 14.5 are artificial. The data for this example have all combinations of the following population densities:

bean — 0, 40,000, 80,000, 120,000, and 160,000 plants per hectare

maize — 20,000, 40,000, and 60,000 plants per hectare

melon — 0, 1000, and 2000 plants per hectare.

TABLE 14.5. Intercrop Yields (kg Grain/ha for Bean and Maize and kg Fruit/ha for Melon) for Three Population Densities (Times 1000) for Maize and Melon and Five for Beans.

Plants/ha		Melon (Plants/ha)							
		0		1			2		
Maize	Bean	Bean	Maize	Bean	Maize	Melon	Bean	Maize	Melon
20	0	—	4934	—	5000	275	—	4500	425
	40	468	3494	550	3494	250	450	3500	400
	80	718	3698	650	3632	225	550	2900	350
	120	775	3632	750	3698	250	700	3500	375
	160	768	4228	740	4228	200	725	4800	325
40	0	—	6446	—	6000	250	—	6200	400
	40	381	5736	400	6000	150	500	6100	280
	80	413	6599	475	5500	175	450	5700	350
	120	663	4660	616	5750	180	600	6500	400
	160	616	6150	500	6100	175	575	6200	375
60	0	—	6485	—	6200	225	—	6000	350
	40	245	7056	250	6870	175	225	6900	400
	80	328	6870	300	7056	150	300	6800	375
	120	323	9124	350	8000	125	250	7500	325
	160	581	5296	450	5000	150	225	4500	300

A mixture of these three crops would appear practical if the beans were planted between the maize rows and the melons (a bush type, e.g.) were planted in the maize rows.

The maize and bean data obtained from Aidar (1978), the third and fourth columns of Table 14.5, are means of three replicate yields from an experiment designed as a randomized complete block. ANOVA tables for bean yields and for maize yields were also presented. From these tables, an error variance for these means (Aidar's error mean square divided by 3, the number of replicates) was obtained. There was 1 missing plot and, hence, the maize error mean square has 27 instead of 28 degrees of freedom. The bean error mean square is associated with 21 degrees of freedom. ANOVAs for the bean, maize, and melon yields in Table 14.5 are given in Table 14.6. The error degrees of freedom for melon, 25, are what would have been had melon been included in the experiment.

For bean yields, significant differences at less than the 5% level were found for bean and maize densities. The bean-by-maize interaction is significant at the 5% level. As shown in Figure 14.5, this interaction is largely due to the low yield for the 60 maize and 120 bean combination. If this combination bean yield had been around 550, there would have been no indication of interaction.

For maize yields, large differences are indicated for maize densities and for the bean-by-maize interaction. The nature of the bean density-by-maize density interaction is depicted in Figure 14.6. The differential yields for the 2 mixtures, 40 maize and 120 beans and 60 maize and 120 beans, account for a large share of

TABLE 14.6. Degrees of Freedom and Mean Squares for Data in Table 14.5.

Source of Variation	Bean d.f.	Bean MS	Maize d.f.	Maize MS	Melon d.f.	Melon MS
Bean = B	3	70,298	4	632,878	4	3,678
Maize = M	2	339,644	2	29,537,314	2	6,520
Melon = L	2	11,449	2	136,449	1	204,118
B × M	6	8,270	8	2,491,944	8	1,903
B × L	6	4,166	8	63,275	4	367
M × L	4	5,895	4	251,812	2	1,750
B × M × L	12	3,551	16	215,380	8	711
Error	21	4,000	27	409,667	25	1,000

Crop in mixture

$F_{.05}(2, 21) = 3.47$ $F_{.05}(2, 27) = 3.35$ $F_{.05}(1, 25) = 4.24$
$F_{.05}(3, 21) = 3.07$ $F_{.05}(4, 27) = 2.73$ $F_{.05}(2, 25) = 3.38$
$F_{.05}(4, 21) = 2.84$ $F_{.05}(8, 27) = 2.30$ $F_{.05}(4.25) = 2.76$
$F_{.05}(6, 21) = 2.57$ $F_{.05}(16, 27) = 2.03$ $F_{.05}(8, 25) = 2.34$
$F_{.05}(12, 21) = 2.25$

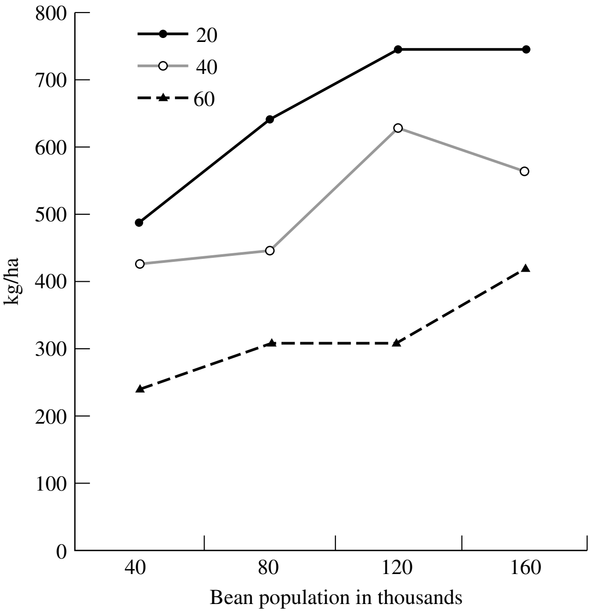

FIGURE 14.5. Population density (plants/ha times 1000) and yields (kg/ha) for bean-by-maize interaction for bean yields.

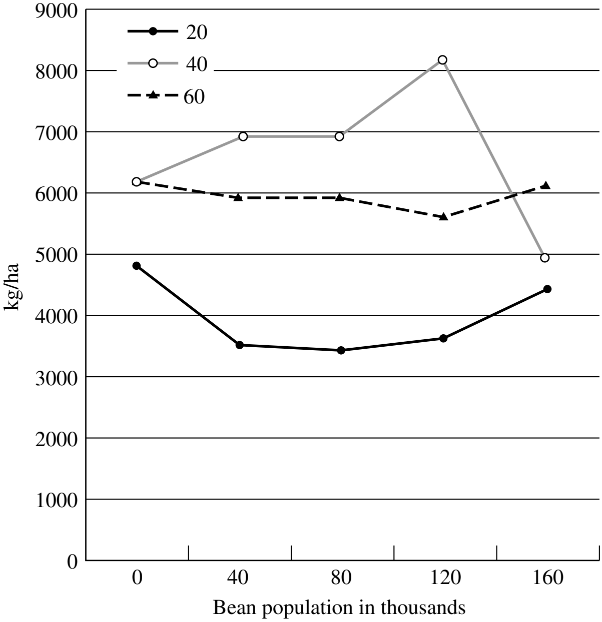

FIGURE 14.6. Population density (plants/ha times 1000) and yields (kg/ha) for bean-by-maize interaction for maize yields.

the interaction sum of squares. Also, the two-mixture yields for 40 maize and 160 beans and 60 maize and 160 beans respond differently than what would be expected on the basis of zero interaction. The trends are different for the maize levels at the three densities. For the 20,000- and 40,000-plant density, the slopes are slightly negative, while the slope is positive for the 60,000 density. An interpretation of these results by the experimenter would need to entail what happened to these particular mixtures in the experiment, as well as to biological theory of why such results are plausible.

For melon yields, the variation in yields between the two melon densities, among the bean densities, and among the maize densities appear real. All the interaction mean squares are nonsignificant.

In addition to the above analyses, the experimenter may wish to model responses as described in the previous section. For an experiment of the sort described here, the following response model would appear appropriate for the sole crop yields. It should be noted that the yield–density curve must go through the point (0, 0), the origin, as zero plants result in zero yields.

$$\bar{y}_{\cdot mi} = \mu + \beta_{1m} d_{mi} + \beta_{2m} d_{mi}^2 + \epsilon_{gd_{mi}(d_{bh} d_{lj})}. \tag{14.30}$$

This equation differs from (14.19) in that the regression function goes through the origin rather than the intercept. Solutions for β_{1m} and β_{2n} are obtained from a solution of the following two equations:

$$\beta_{1m}\,\Sigma\, d_{mi}^2 + \beta_{2m}\,\Sigma\, d_{mi}^3 = \Sigma\, d_{mi}\,\bar{y}_{\cdot mi}, \tag{14.31}$$

$$\beta_{1m}\,\Sigma\, d_{mi}^3 + \beta_{2m}\,\Sigma\, d_{mi}^4 = \Sigma\, d_{mi}^2\,\bar{y}_{\cdot mi}. \tag{14.32}$$

For this example,

$$0 + 2 + 4 + 6 = \Sigma\, d_{mi},$$

$$0^2 + 2^2 + 4^2 + 6^2 = 56 = \Sigma\, d_{mi}^2,$$

$$0^3 + 2^3 + 4^3 + 6^3 = 288 = \Sigma\, d_{mi}^3,$$

$$0^4 + 2^4 + 4^4 + 6^4 = 1568 = \Sigma\, d_{mi}^4,$$

$$2(4934) + 4(6446) + 6(6485) = 74,562 = \Sigma\, d_{mi}\,\bar{y}_{\cdot mi},$$

$$4(4934) + 16(6446) + 36(6485) = 356,332 = \Sigma\, d_{mi}^2\,\bar{y}_{\cdot mi}.$$

With these results, equations (14.31) and (14.32) become

$$56\beta_{m1} + 288\beta_{m2} = 74,562,$$

$$288\beta_{m1} + 1568\beta_{m2} = 356,332.$$

Solution of the above results in $\beta_{m1} = 2938$ and $\beta_{m2} = -312.35$. The predicted yields on the basis of this regression function is

$$\hat{Y}_{mi} = 2938 d_{mi} - 312.35 d_{mi}^2,$$

which for $d_{mi} = 2$ (20,000 density) is

$$\hat{Y}_{m2} = 2(2938) - 4(312.35) = 4627.$$

The yields $\bar{y}_{\cdot mi}$ and the predicted yields $\hat{Y}_{mi}$ are plotted in Figure 14.7. From the curve, for $\hat{Y}_{mi}$, the number of plants per hectare for optimum yield is computed as $-\hat{\beta}_1/2\hat{\beta}_2 = 47,028$ plants per hectare.

Using equations (14.24) and (14.30), the $\hat{\gamma}_{m(bl)}(hij)$ effects are computed as follows. Let $h = 20$, $i = 40$, and $j = 0$; then (see Table 14.7)

$$\hat{\gamma}_{m(bl)}(20, 40, 0) = 3494 - 2(2938) + 4(312.35) = -1133.$$

Also,

$$\hat{\gamma}_{m(bl)}(20, 80, 0) = 3698 - 4627 = -929,$$

$$\vdots$$

$$\hat{\gamma}_{m(bl)}(60, 160, 0) = 5296 - 6383 = -1088,$$

$$\hat{\gamma}_{m(bl)}(20, 40, 1) = 3494 - 4627 = -1133,$$

$$\vdots$$

$$\hat{\gamma}_{m(bl)}(60, 160, 2) = 4500 - 6383 = -1883.$$

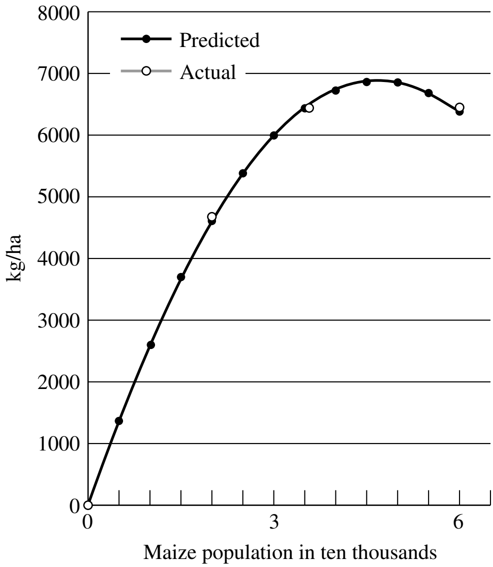

FIGURE 14.7. Sole crop maize yields (kg/ha) for different maize densities per hectare.

TABLE 14.7. Mixture Density Effects for Maize Yields, $\hat{\gamma}_{m(bl)}(hij)$, from Equation (14.24) for Densities $h = 20, 40, 60$, $i = 0, 40, 80, 120, 160$, and $j = 0, 1, 2$ (times 1000).

$h, i, 0$	$\hat{\gamma}_{m(bl)}(hi0)$	$h, i, 1$	$\hat{\gamma}_{m(bl)}(hi1)$	$h, i, 2$	$\hat{\gamma}_{m(bl)}(hi2)$
20, 40, 0	−1133	20, 40, 1	−1133	20, 40, 2	−1127
20, 80, 0	−929	20, 80, 1	−995	20, 80, 2	−1727
20, 120, 0	−995	20, 120, 1	−929	20, 120, 2	−1127
20, 160, 0	−399	20, 160, 1	−399	20, 160, 2	173
40, 40, 0	−1018	40, 40, 1	−754	40, 40, 2	−654
40, 80, 0	−155	40, 80, 1	−1254	40, 80, 2	−1054
40, 120, 0	−2094	40, 120, 1	−1004	40, 120, 2	−254
40, 160, 0	−604	40, 160, 1	−654	40, 160, 2	−554
60, 40, 0	673	60, 40, 1	487	60, 40, 2	517
60, 80, 0	487	60, 80, 1	673	60, 80, 2	417
60, 120, 0	2741	60, 120, 1	1617	60, 120, 2	1117
60, 160, 0	−1087	60, 160, 1	−1383	60, 160, 2	−1883

The three largest mixture–density effects are for the mixtures (40, 120, 0), (60, 120, 0), and (60, 160, 2). Other large effects are for mixtures (60,120,1) and (20, 80, 2). A variance for a mixture–density effect is computed as for density d_{i^*}:

$$\mathrm{V}(\hat{\gamma}_{m(bl)}(hij) = \sigma_\epsilon^2/r + \mathrm{Var}\left(\beta_{1m}d_{i^*} + \beta_{2m}d_{i^*}^2\right)$$

$$= \frac{\sigma_\epsilon^2}{r}\left[1 + \left\{\left(\Sigma\, d_i \left(d_{i'*}\Sigma\, d_i^4 - d_{i*}^2 \Sigma\, d_i^3\right)\right.\right.\right.$$
$$+ \; \Sigma\, d_i^2 \left(d_{i*}^2 \Sigma\, d_i^2 - d_{i*} \Sigma\, d_i^3\right)\Big)\Big/$$
$$\left.\left.\left(\Sigma\, d_i^2 \Sigma\, d_i^4 - \left(\Sigma\, d_i^3\right)^2\right)\right\}^2\right],$$

which for the above example for $d_{i*} = 2$ is

$$409,667(1 + 12[2(1568) - 2^2(288)] + 56[2^2(56) - 2(288)])$$
$$= \; 409,667(1 + 0.70914) = 700,178.$$

For Densities 4 and 6, the variances are

$$409,667(1 + 1.3407) = 958,908$$

and

$$409,667(1 + 0.89751) = 777,347.$$

The standard errors for $d_{i*} = 2, 4$, and 6 are 837, 979, and 882, respectively. A 5% significant difference is $t_{.05}(27) = 2.052$ times the standard error for each $d_{i*} = 2$, 4, and 6; these are 1717, 2009, and 1810, respectively. The mixture–density effects which exceed these significant differences are for mixtures (40, 120, 0), (60, 120, 0), (20, 80, 2), and (60, 160, 2).

14.8 Analysis for c_i Lines for Crop i

In place of using a single line for each of the crops in the intercrop mixture, an experimenter may wish to use several lines of one or more of the crops. Also, several lines for one, for two, ..., or for all crops may be included in an experiment. For l_i lines of crops i, there will be a total of $l_1 \times l_2 \times \ldots \times l_c = L$ line combinations. As is obvious, L can quickly become large. If one of the five designs is included for each of the L combinations, a large number of experimental units and combinations result. Hence, methods should be considered which reduce the size of an experiment to a manageable size. This means that the lines to be included should be carefully selected in order to make L as small as possible.

For Designs 1 and 5 in Section 14.2, it may be possible to have the experimental unit for combination $d_{ch}d_{mi}d_{bj}$, say, large enough to accommodate the L line combinations. For example, if $L = 8$, the experimental unit for level combination $d_{ch}d_{mi}d_{bj}$ could be divided into $L = 8$ split plot experimental units to accommodate the $L = 8$ line combinations. Using the usual randomization procedure for the experiment design, an ANOVA for this type of design is given in Table 14.8. By partitioning the degrees of freedom into single degree of freedom contrasts, several of the degree of freedom contrasts should be estimates of the error variance for such an experiment. If so, the Daniel (1959) and Krane (1963) procedures may be used to estimate error variances and construct confidence intervals.

TABLE 14.8. ANOVA Partitioning of Degrees of Freedom for $L = 8$ Line Combinations as Split Plots of the Fractional Replicate of Figure 14.4.

Source of variation	d.f.	SS	MS
Total	$19L = 19(8) = 152$		
Correction for mean	1		
Treatment combinations	$(v - 1) = 18$		
Crop one = A	2		
Crop two = B	2		
Crop three = C	2		
A × B	4		
A × C	4		
B × C	4		
Split Plot	$19(8 - 1) = 133$		
Line combination = LC	7		
LC × A	14		
LC × B	14		
LC × C	14		
LC × A × B	28		
LC × A × C	28		
LC × B × C	28		

As an alternative to the above, the line combinations may be used as whole plots for the density level combinations as the split plots. Then simply interchange the roles of the line and density combinations as shown in Table 14.9 to obtain this partitioning of degrees of freedom. Again, the appropriate single degree of freedom contrasts should be made in order to have a complete analysis for the experiment.

For Designs 2 and 3, it would appear that the line combinations should be the whole plot treatments. Following the ideas in Tables 14.8 and 14.9, the partitioning of the degrees of freedom and sums of squares is straightforward. Fractional replicates of all possible line combinations and of all possible density level combinations may be used if desired. ANOVAs similar to the above may be used. Also, analyses for each crop as well as for the combination of crop yields may be used.

14.9 Summary and Discussion

Five treatment designs are presented for studying yield–density relationships for mixtures of three or more crops. There are many more that could be considered, but the ones given should enable the reader to develop others appropriate for the situation in question. Likewise, response model equations are presented for monocrops and mixtures. Models using the lowest density levels of all crops but the one under consideration are also presented. As pointed out, other and/or more complex models for yield–density relationships may be used where appropriate.

TABLE 14.9. ANOVA Partitioning of Degrees of Freedom for $L = 8$ Whole Plots and with the 19 Density Combinations of Figure 14.4 as Split Plots.

Source of variation	d.f.	SS	MS
Total	8(19) = 152		
Correction for mean	1		
Line combination = LC	7		
Split Plots	144		
Crop one = A	2		
Crop two = B	2		
Crop three = C	2		
A × B	4		
A × C	4		
B × C	4		
A × LC	14		
B × LC	14		
C × LC	14		
A × B × LC	28		
A × C × LC	28		
B × C × LC	28		

Three of the proposed designs (2, 3, and 4) are parsimonious in that a wide range of density levels may be studied using a small amount of experimental space and material. These parsimonius designs allow studies on wide ranges of densities and numbers of crops within a doable framework.

Designs involving replacement of one crop by another (replacement series) have not been discussed in this chapter. This discussion has been relegated to Chapter 17 in the same manner as it was done in Volume I. There, this topic is discussed in Chapter 8. Several parsimonious designs for studying response–density relations are presented in Chapter 17.

Since many analyses are often required to extract the information from the data obtained in an experiment, it is recommended that statistical analyses like those in this chapter be done; that is, analyses are made for each crop in a mixture as well as for each of the created variables (e.g., relative land equivalent ratio, relative system value, etc.). Single statistical analyses, which are most frequently used with monocultures, are usually insufficient for extracting the information in an intercropping experiment.

14.10 Problems

14.1 Assume that the sole crop bean yields are 400, 500, 600, 800, and 700 kg/ha for densities 0, 40,000, 80,000, 120,000, and 160,000 plants/ha. Obtain solutions for the linear and quadratic coefficients and prepare a table similar to Table 14.7 for mixture–density effects for bean yields.

14.2 Run an analysis of variance on the $\hat{\gamma}_{b(ml)}(hij)$ values computed in Problem 14.1. Do likewise for the data in Table 14.7 for the $\hat{\gamma}_{m(bl)}$ effects. Interpret the results.

14.3 Propose alternate models for mixture–density effects and determine under what circumstances your models would be appropriate.

14.4 Compare the linear regression coefficients for the $\hat{\gamma}_{m(bl)}$ effects for each of the three maize densities.

14.5 Compute variances for the $\hat{\gamma}_{b(ml)}(hij)$ effects in Problem 14.1 and determine which, if any, of the effects exceed $t_{.05}(21)$ times the standard error of an effect.

14.11 Literature Cited

Aidar, H. (1978). Estudo populações de plantas em dois sistemas de culturas associadas de milho e feijão. Doctor Scientiae, Universidade Federal de Viçosa.

Anderson, D.A. and W.T. Federer (1975). Possible determinant values for square (0,1)-matrices in fractional replication. *Utilit. Math.* **7**, 135-150.

Anderson, D.A. and W.T. Federer (1995). Representation and bounds for the general fractional factorial. *Commun. in Statist.: Theory Applic.* **24**, 363-373.

Birnbaum, A. (1959). On the analysis of factorial experiments without replication. *Technometrics* **1**, 343-357.

Daniel, C. (1959). Use of half-normal plots in interpreting factorial two-level experiments. *Technometrics* **1**, 311-341.

Ezumah, H.C., W.T. Federer, H.G. Gauch, Jr., and R.W. Zobel (1991). Analysis and interpretation of the maize component of maize and cassava systems with nitrogen in a multilocation experiment. BU-1129-M in the Technical Report Series of the Biometrics Unit, Cornell University, Ithaca, NY.

Federer, W.T. and B.T. Scully (1993). A parsimonious statistical design and breeding procedure for evaluating and selecting desirable characteristics over environments. *Theoret. Appl. Genet.* **86**, 612-620.

Gauch, H.G. (1988). Model selection and validation for yield trials with interaction. *Biometrics* **44**, 705-715.

Gauch, H.G. and R.W. Zobel (1988). Accuracy and selection success in yield trial analyses. *Theoret. Appl. Genet.* **77**, 473-481.

Grimes, B.A. and W.T. Federer (1984). Comparison of means from populations with unequal variances. In *W.G. Cochran's Impact on Statistics*, (Editors: P.S.R.S. Rao and J. Sedransk), John Wiley and Sons, Inc., New York, pp. 353-374.

Krane, S.A. (1963). Half-normal plots for multi-level factorial experiments. *Proc. Eighth Conference Design of Experiments in Army Research Development and Testing*, ARO-D Report 63-2, U.S. Army Research Office, Durham, NC, pp. 261-285.

Morales, M.F.M. (1993). Estimation and design of yield–density models for Intercropping Experiments. Ph.D. dissertation, University of California, Riverside.

Raktoe, B.L., A. Hedayat, and W.T. Federer (1981). *Factorial Designs*. John Wiley and Sons, Inc., New York.

CHAPTER 15

Mixing Ability Effects When Individual Cultivar Responses Are Available

15.1 Introduction

In Volume I, Chapter 6, biological models for cultivars in mixtures were presented for m cultivars in mixtures of $n = 2$ cultivars. In this chapter, we generalize the concepts and statistical methods to m cultivars in mixtures of size n cultivars following the approach of Federer and Raghavarao (1987). These authors considered minimal designs to estimate all types of mixing abilities for m cultivars in mixtures of size n. They built upon the ideas in Hall (1976), Federer (1979), and Federer *et al.* (1976) to obtain the theoretical results for this chapter.

Knowing and understanding the type and nature of biological interactions of cultivars in mixtures can be of considerable practical usefulness in obtaining mixtures for general usage by farmers. In other types of mixtures such as drugs, such information as may be obtained from the statistical procedures of this chapter can be extremely valuable, and perhaps vital, for prescribing mixtures of drugs for patients.

Types and nature of various effects for a mixture are discussed in Section 15.2. Minimal treatment designs and response model equations for various types of effects are presented in Section 15.3. Designs and response models are presented first for general mixing ability effects, GMA. This is followed by a presentation of designs and models for GMA and bi-specific mixing effects, BSMA. Then, designs and models are discussed for the situation where GMA, BSMA, and tri-specific mixing effects, TSMA, are desired. Also, designs and models for GMA, BSMA, TSMA, and quat-specific mixing effects, QSMA, are discussed. The case for the kth specific mixing effect estimation ends Section 15.3. In Section 15.4, solutions and variances are given for estimated effects for each of the situations in Section 15.3. As the number of cultivars, n, in a mixtures increases, the complexity of the algebraic expressions also increases. Owing to this, the algebraic developments were relegated to Appendix 15.1. Also, owing to the algebraic complexity, use was made of computer software programs such as MAPLE and MATHEMATICA, to obtain solutions, as described in Appendix 15.2. The increase in algebraic complexity for larger n can be compensated for to some extent by using this software. Mixtures $n = 2, 3$, and 4 are detailed herein. By extending the computer programs in Appendix 15.2, solutions and variances of differences of effects may be obtained for $n = 5, 6, \ldots$ for specific values of m.

Four numerical examples are presented to demonstrate the application of the formulas presented and to demonstrate another analysis for the examples of Chapters 12 and 13. Again, it is emphasized that several statistical analyses may be required to glean the information from an intercropping experiment. Problems are presented for the reader to apply the results of this chapter and to extend the concepts.

15.2 Type and Nature of Effects

The various biological types of effects considered in this chapter are those related to a cultivar's ability to mix either specifically or generally with $n-1$ other cultivars. A cultivar may perform well in intercrop mixtures of size n with all other $m - 1$ cultivars under consideration. A cultivar may do especially well or poorly with a particular mixture of n cultivars. The former type is known as general mixing ability and the latter as specific mixing ability. These concepts have been discussed in Hall (1976), Federer (1979), Federer *et al.* (1976), and Volume I, Chapters 6 and 7.

15.2.1 General Mixing Ability

A general mixing ability (GMA) effect is a measure of how well or how poorly a particular cultivar performs in mixtures of size n of m cultivars. The measure is in reference to this particular set of m cultivars under consideration in an experiment involving mixtures of n cultivars. If a cultivar performs well with the other $m - 1$ cultivars in mixtures of size n, then it is said to have a high GMA. Such a cultivar is desirable. If a cultivar performs poorly in all mixtures of size n in which it is a member, the cultivar is undesirable for intercropping purposes.

A cultivar effect is the sum of its relative performance as a sole crop, say τ_i, plus its performance in a mixture, say δ_i. Thus, a cultivar effect is denoted as $\tau_i + \delta_i$, and if only mixtures of size n are included, the individual solutions for the terms τ_i and δ_i cannot be obtained. The δ_i refers to the GMA effect, but it should be made clear that a cultivar effect of $\tau_i + \delta_i$ is all that may be obtained. This should be clearly understood when interpreting cultivar and GMA effects as they are different.

15.2.2 Bi-specific Mixing Ability

A bi-specific mixing ability (BSMA) effect refers to a first-order or two-factor interaction effect of a pair of cultivars in mixtures of size n. Two cultivars may perform beneficially or detrimentally when they appear together in a mixture. A positive value for a BSMA effect is desirable, as it complements the GMA effects for the pair of cultivars since the overall effect of a cultivar is the sum of its individual effects. BSMA effects can vary with the size of a mixture (see Hall, 1976). If an experimenter is interested only in GMA and BSMA effects, then n may be taken as 2. If BSMA effects vary with n, then it will be necessary to vary n in the experiment to assess BSMA effects for various n.

As opposed to ordinary interaction terms, a BSMA effect is obtained for each i and each j in the pair ij. A specific combining ability effect in genetic experiments is obtained only for the pair ij and not for i alone and j alone. We denote these two interaction terms as $\lambda_{i(j)}$ and $\lambda_{j(i)}$, where the symbol in parentheses means "in the presence of." $\lambda_{i(j)}$ is the contribution of cultivar i to the interaction of cultivar i with j, i.e., $\lambda_{i(j)} + \lambda_{(i)j} = \lambda_{ij}$. These effects have solutions when responses are obtained for each member of the pair ij.

15.2.3 Tri-specific Mixing Ability

A tri-specific mixing ability (TSMA) effect is a second-order or three-factor, interaction effect of three cultivars in a mixture of size n cultivars. This effect may vary with the number n of intercrops in a mixture. If so, then n, the number of cultivars in a mixture, may need to be varied in an experiment. A trio of cultivars may mix well or poorly in a mixture. A TSMA effect may be independent of GMA and BSMA effects. A positive-valued TSMA effect is desirable, which indicates that the specific trio performs well in a mixture.

In an ordinary interaction effect, only one value is obtained, but when individual responses from each member of a trio are available, a TSMA effect is obtainable for each member of a trio. If an ordinary interaction effect is denoted as π_{ijk}, say, then the three corresponding TSMA effects are denoted as $\pi_{i(j,k)}$, $\pi_{j(i,k)}$, and $\pi_{k(i,j)}$, where the symbols in parentheses mean "in the presence of." Knowing the individual interaction components provides additional information over the ordinary interaction term π_{ijk} which is the sum of the components, i.e., $\pi_{i(jk)} + \pi_{j(ik)} + \pi_{k(ij)} = \pi_{ijk}$.

15.2.4 k-specific Mixing Effect

A $(k - 1)$th-order or k-factor interaction effect of a n-tuplet of cultivars is denoted as a kth specific mixing ability (KSMA) effect. A four-cultivar mixing effect is denoted as quat-specific mixing ability (QSMA) effect. A five-cultivar specific mixing ability effect is denoted as quint-specific mixing ability (QuSMA) effect. In general, a k-cultivar specific mixing ability effect is denoted as a k-specific mixing ability (KSMA) effect. When responses for individual cultivars are available, a KSMA effect will be obtained for each member of the k-tuple, in the mixture of n cultivars, $k \leq n$. As k increases, it would be expected that the size of the specific mixing ability effects would decrease. However, as pointed out in Chapter 11, surprising results from intercropping experiments are ever-present. Although the likelihood may be small, four-, five-, six-, and even higher-factor interaction effects are possible. The complexity of intercropping systems will be related to the existence of these higher-order interactions. In an intercropping experiment, Ezumah *et al.* (1991) found significant two-, three-, and four-factor interactions in a four-factor experiment which were meaningful and interpretable. If more factors had been included, it is possible that higher-order interactions would have been present. Significant and meaningful five-factor interactions have been encountered by the author. The rareness of such interaction should not be interpreted as nonexistence. Their presence increases as the complexity of a biological system increases. Intercropping systems often involve complex biological relationships.

15.3 Minimal Designs and Response Models

Hall (1976), Federer *et al.* (1976), and Federer and Raghavarao (1987) have discussed minimal designs for estimating various types of effects in intercropping experiments. We have been and will continue to use the notation in the last reference above. Minimal designs will vary depending upon whether only GMA, GMA plus BSMA, GMA plus BSMA plus TSMA, etc. effects are desired. The treatment design used must be such that solutions for the parameters involved are possible. First, we consider minimal designs for obtaining solutions for only GMA effect parameters. Then minimal designs for obtaining solutions for GMA and BSMA effect parameters will be discussed. This is followed by a discussion of treatment designs for GMA + BSMA + TSMA effects, and for GMA + BSMA + TSMA + QSMA effects.

The last section is a discussion of treatment designs for GMA + BSMA + $\cdots$ + KSMA effects. Also, the designs discussed are equal-sizes mixtures n for all mixtures denoted as S_α, $\alpha = 1, \cdots, v$ the number of mixtures. Comments on designs for variable sizes for n are presented.

15.3.1 Minimal Designs for Cultivar and GMA Effects

A minimal treatment design for obtaining estimates of GMA effects only is to include each of the m cultivars as sole crops and all m cultivars in a single mixture of $n = m$. For various reasons such as variance differences and precision, it would be wise to include the mixture of m cultivars more than once in each complete block. This allows the estimation of variances for each GMA effect. For the example presented below, the $m = 6$ cultivar mixture was included once. Denoting the sole crop mean as $\mu_h + \tau_h$, the crop mean in mixtures is denoted as $(\mu_h + \tau_h + \delta_h)/n$, where δ_h is the GMA effect for crop h. Here, δ_h is the GMA effect and $\tau_h + \delta_h$

is the cultivar h effect. It is assumed that interaction effects are nonexistent. The factor $1/n$ is needed to put these effects on the same basis as sole crop yields.

Response model equations for mixtures of size $n = 3$ cultivars h, i, and j, say, when the treatments are in a randomized complete block design (RCBD) with mixtures $S_1, S_2, \ldots, S_v$ as the experimental units and the v treatments with each complete block are

$$Y_{gh(i,j)} = \big(\mu_h + \rho_{gh} + (\tau_h + \delta_h)\big)\big/ 3 + \epsilon_{gh(i,j)}, \tag{15.1}$$

$$Y_{gi(h,j)} = \big(\mu_i + \rho_{gi} + (\tau_i + \delta_i)\big)\big/ 3 + \epsilon_{gi(h,j)}, \tag{15.2}$$

$$Y_{gj(h,i)} = \big(\mu_j + \rho_{gj} + (\tau_j + \delta_j)\big)\big/ 3 + \epsilon_{gj(h,i)}, \tag{15.3}$$

where $\mu_h + (\tau_h + \delta_h)$ is the mean of cultivar h summed over all combinations of cultivars i and j for $h \neq i, h \neq j, i \neq j$, τ_h is the cultivar effect as a sole crop, δ_h is the GMA effect parameter, and $\epsilon_{gh(i,j)}$ is a random error effect distributed with mean zero and variance $\sigma^2_{\epsilon h}$. $Y_{gh(i,j)}$ is the response for cultivar h in the mixture hij in the gth complete block. The symbols in (15.2) and (15.3) are defined similarly for cultivar i and j responses. In certain situations, $\mu_h = \mu_i = \mu_j = \mu$. This is a form of the linear model discussed by Federer and Raghavarao (1987). Whether or not this situation is tenable will be determined by the nature of the m cultivars in the experiment and the goals of the experimenter. It should be noted that solutions are possible for $\mu_x + \tau_x + \delta_x$, $x = i, j, k$, but not for the individual terms. If $\mu_x = \mu$, the solutions are possible for μ and $\tau_x + \delta_x$ when a restriction such as $\sum_{x=1}^{m}(\tau_x + \delta_x) = 0$ is imposed. In order to obtain solutions for δ_x, solutions for μ and τ_x or $\mu_x + \tau_x$, will need to be obtained from sole crop responses under the restriction $\sum_{x=1}^{m} \tau_x = 0$. The restriction that $\sum_{x=1}^{m} \delta_x = 0$ is not tenable as all δ_x could be negative, all could be positive, or there could be a mixture of positive and negative values for GMA effects. When interpreting whatever solution is obtained, careful thought of the meaning of any solution is required in order to make correct interpretations and inferences.

15.3.2 Minimal Designs for Cultivar Plus BSMA Effects

In order to obtain solutions for cultivar effects, it was necessary to have at least m mixtures of size $n = 2$ for the m cultivars. To obtain solutions for both cultivar and BSMA effects for m cultivars, at least $rm(m-1)$ responses from the v mixtures must be available for response model equations of the following form for an RCBD-designed experiment:

$$Y_{gh(i)} = \big(\mu_h + \rho_{gh} + (\tau_h + \delta_h)\big)\big/ 2 + \beta_{h(i)} + \epsilon_{gh(i)}, \tag{15.4}$$

$$Y_{g(h)i} = \big(\mu_i + \rho_{gi} + (\tau_i + \delta_i)\big)\big/ 2 + \beta_{(h)i} + \epsilon_{g(h)i}, \tag{15.5}$$

where $Y_{gh(i)}$ is the response for cultivar h in the pair hi (in the presence of cultivar i) in the gth replicate, $g = 1, \ldots, r$, $h = 1, \ldots, m$, $i = 1, \ldots, m$, $i \neq j$, $\mu_h + \tau_h + \delta_h$ is the mean effect of cultivar h in the mixture and similarly for cultivar i as described in the previous section, $\beta_{h(i)}$ is the BSMA effect for cultivar

h when paired with cultivar i, $\beta_{(h)i}$ is the BSMA effect for cultivar i when paired with cultivar h, and $\epsilon_{gh(i)}$ and $\epsilon_{g(h)i}$ are random error effects distributed with mean zero and variances $\sigma^2_{\epsilon h}$ and $\sigma^2_{\epsilon i}$, respectively. The remarks in the previous subsection about μ_h, τ_h, τ_i, δ_h, and δ_i apply here as well.

A minimal treatment design for obtaining solutions for the parameters of response equations (15.4) and (15.5) for $n = 2$ is the irreducible balanced incomplete block design for all possible combinations of m cultivars taken two at a time, or $m!/2!(m-2)! = m(m-1)/2$ combinations. For $m = 4$ and $n = 2$, the $m(m-1)/2$ mixtures or blocks of the BIBD (balanced incomplete block design) are $S_1 = (1, 2)$, $S_2 = (1, 3)$, $S_3 = (1, 4)$, $S_4 = (2, 3)$, $S_5 = (2, 4)$, and $S_6 = (3, 4)$. This is a BIBD with $v = 4$, $k = 2$, $b = 6$, $r = 3$, and $\lambda = 1 =$ number of times cultivar h is paired with cultivar i in the b blocks (mixtures). The number of cultivars, m, must be greater than two in order for BSMA effect to have meaning.

Since there must be at least $m(m - 1)$ responses in order to obtain solutions for $(\mu_h + \tau_h + \rho_h)$, $\beta_{h(i)}$, and $\beta_{(h)i}$ parameters, it is possible to do this by using m mixtures of size $n = m - 1$ (i.e., a BIBD with $v = b = m$, $k = r = m - 1$, $\lambda = m - 2$) in a BIBD arrangement. Using this treatment design would require that n, the size of the mixture, would not affect the BSMA effect; that is, this effect would be the same regardless of whether mixtures of two cultivars or mixtures of $m - 1$ cultivars were used. For many studies, it would appear that mixtures of $n = 2$ would be preferable, although the parsimony achieved with mixtures of $m - 1$ cultivars is appealing even though their precision is less.

15.3.3 Minimal Designs for Cultivar Plus BSMA Plus TSMA Effects

Treatment designs for obtaining solutions for cultivar effects, BSMA effects, and TSMA effects must have mixtures of $n \geq 3$ cultivars in a mixture. For $n = 3$, response model equations for cultivars h, i, and j in the mixture are

$$Y_{gh(i,j)} = \frac{1}{3}\left(\mu_h + \rho_{gh} + (\tau_h + \delta_h)\right) + \frac{2}{3}(\beta_{h(i)} + \beta_{h(j)}), \\ + \pi_{h(i,j)} + \epsilon_{gh(i,j)}, \tag{15.6}$$

$$Y_{gi(h,j)} = \frac{1}{3}\left(\mu_i + \rho_{gi} + (\tau_i + \delta_i)\right) + \frac{2}{3}(\beta_{i(h)} + \beta_{i(j)}) \\ + \pi_{i(h,j)} + \epsilon_{gi(h,j)}, \tag{15.7}$$

$$Y_{gj(h,i)} = \frac{1}{3}\left(\mu_j + \rho_{gj} + (\tau_j + \delta_j)\right) + \frac{2}{3}(\beta_{j(h)} + \beta_{j(i)}) \\ + \pi_{j(h,i)} + \epsilon_{gj(h,i)}, \tag{15.8}$$

where $Y_{gh(i,j)}$ is the response for cultivar h, $h = 1, \ldots, m$, in mixture hij in the gth replicate, $\pi_{h(i,j)}$ is the TSMA effect for cultivar h for the trio hij (or in the presence of cultivars i and j), and similarly for $Y_{gi(h,j)}$, $Y_{gj(h,i)}$, and other terms for cultivars i and j. The remaining parameters are as described in the previous two subsections. The same equations would be used for the triplet of cultivars h, i, and j from treatment designs where $n > 3$. The number of responses used for an effect is determined by the number of factors in the interaction term.

A minimal design for obtaining solutions for the parameters in equations (15.6) to (15.8) is the irreducible BIBD for $m \geq 2t$ (t=order of interaction) for m even and $m \geq 2t + 1$ for m odd in blocks of size $n = 3$. This design has all possible combinations of m items taken three at a time. Thus, the number of blocks of the BIBD, or mixtures, is $m!/3!(m-3)! = m(m-1)(m-2)/6 = b = v$. There will be $\lambda = m - 2$ occurrences of pairs of cultivars in the b blocks. Any particular cultivar will occur $r = (m-1)(m-2)/2$ times in the b mixtures. Each triplet of cultivars will occur once in the b blocks. To illustrate, the design for $m = 6$, $v = 20$, and $n = 3$ is $S_1 = (1, 2, 3)$, $S_2 = (1, 2, 4)$, $S_3 = (1, 2, 5)$, $S_4 = (1, 2, 6)$, $S_5 = (1, 3, 4)$, $S_6 = (1, 3, 5)$, $S_7 = (1, 3, 6)$, $S_8 = (1, 4, 5)$, $S - 9 = (1, 4, 6)$, $S_{10} = (1, 5, 6)$, $S_{11} = (2, 3, 4)$, $S_{12} = (2, 3, 5)$, $S_{13} = (2, 3, 6)$, $S_{14} = (2, 4, 5)$, $S_{15} = (2, 4, 6)$, $S_{16} = (2, 5, 6)$, $S_{17} = (3, 4, 5)$, $S_{18} = (3, 4, 6)$, $S_{19} = (3, 5, 6)$, and $S_{20} = (4, 5, 6)$. Here again, as with the two previous cases, in order to obtain solutions for the GMA effects, or δ_h parameters, sole crops need to be included in the experiment. Also, if $\mu_h = \mu_i = \mu_j = \mu$, then solutions for $\tau_h + \delta_h$, $\tau_i + \delta_i$, and $\tau_j + \delta_j$ are possible under the restriction $\sum_{x=1}^{m}(\tau_x + \delta_x) = 0$. Also, use may be made of BIBDs balanced for occurrence of triplets (not just pairs) of cultivars with $n > 3$ cultivars in a mixture. Whether or not an experimenter would want to do this would depend on n, the size of the mixture, to be used in practice.

15.3.4 *Minimal Designs for Cultivar Plus BSMA Plus TSMA Plus QSMA Effects*

Response model equations for $n = 4$ cultivars in a mixture when cultivar effects, bi-specific mixing effects (BSMA), tri-specific mixing effects (TSMA), and quat-specific mixing effects (QSMA) are present are

$$\begin{aligned} Y_{gh(i,j,k)} = {} & \frac{1}{4}\left(\mu_h + \rho_{gh} + (\tau_h + \delta_h)\right) + \frac{2}{4}(\beta_{h(i)} + \beta_{h(j)} + \beta_{h(k)}) \\ & + \frac{3}{4}(\pi_{h(i,j)} + \pi_{h(i,k)} + \pi_{h(j,k)}) + \gamma_{h(i,j,k)} + \epsilon_{gh(i,j,k)}, \end{aligned} \tag{15.9}$$

$$\begin{aligned} Y_{gi(h,j,k)} = {} & \frac{1}{4}\left(\mu_i + \rho_{gi} + (\tau_i + \delta_i)\right) + \frac{2}{4}(\beta_{i(h)} + \beta_{i(j)} + \beta_{i(k)}) \\ & + \frac{3}{4}(\pi_{i(h,j)} + \pi_{i(h,k)} + \pi_{i(j,k)}) + \gamma_{i(h,j,k)} + \epsilon_{gi(h,j,k)}, \end{aligned} \tag{15.10}$$

$$\begin{aligned} Y_{gj(h,i,k)} = {} & \frac{1}{4}\left(\mu_j + \rho_{gj} + (\tau_j + \delta_j)\right) + \frac{2}{4}(\beta_{j(h)} + \beta_{j(i)} + \beta_{j(k)}) \\ & + \frac{3}{4}(\pi_{j(h,i)} + \pi_{j(h,k)} + \pi_{j(i,k)}) + \gamma_{j(h,i,k)} + \epsilon_{gj(h,i,k)}, \end{aligned} \tag{15.11}$$

$$Y_{gk(h,i,j)} = \frac{1}{4}\left(\mu_k + \rho_{gk} + (\tau_k + \delta_k)\right) + \frac{2}{4}(\beta_{k(h)} + \beta_{k(i)} + \beta_{k(j)}) + \frac{3}{4}(\pi_{k(h,i)} + \pi_{k(h,j)} + \pi_{k(i,j)}) + \gamma_{k(h,i,j)} + \epsilon_{gk(h,i,j)}, \quad (15.12)$$

where $Y_{gh(i,j,k)}$ is the response for cultivar h from the mixture $hijk$ in the gth replicate, μ_h is a general mean effect for cultivar h, ρ_{gh} is the gth block effect for response for cultivar h, $(\tau_h + \delta_h)$ is the cultivar effect for h, $\beta_{h(i)}$, $\beta_{h(j)}$, and $\beta_{h(k)}$ are BSMA effects for h in the presence of cultivars i, j, and k, respectively, the factor $\frac{2}{4}$ is needed as only one-half of the experimental unit is occupied by this pair of cultivars, $\pi_{h(i,j)}$, $\pi_{h(i,k)}$, and $\pi_{h(j,k)}$ are TSMA effects for h in the presence of cultivars i and j, i and k, and j and h, respectively, the factor $\frac{3}{4}$ is used as only three-fourths of the experimental unit is occupied by these cultivars, $\gamma_{h(i,j,k)}$ is the QSMA for h in the presence of i, j, and k, and $\epsilon_{gh(i,j,k)}$ is a random error effect for cultivar h responses distributed with mean zero and variance $\sigma^2_{\epsilon h}$. The symbols in (15.10), (15.11), and (15.12) are defined similarly. The QMSA are four-factor interaction effects for *each* cultivar rather than the usual four-factor interaction effect which is the sum of individual cultivar interaction effects.

In order to obtain solutions for all parameters in the response models (15.9) to (15.12), it is necessary that $n \geq 4$ cultivars and $m \geq 8$ cultivars in a mixture (see Tables 15.1 and 15.2); that is for a tth-order, or $(t+1)$st-factor interaction, $m \geq 2t$ and $n \geq t+1$ (see Federer and Raghavarao, 1987). A minimal design for solution for the effects in equations (15.9) to (15.12) for m cultivars in mixtures of size $n = 4$ is all possible combinations of m items taken $n = 4$ at a time, or $m!/4!(m-4)! = m(m-1)(m-2)(m-3)/24$ combinations. For $m = 8$ and $n = 4$, the combinations are $S_1 = (1, 2, 3, 4)$, $S_2 = (1, 2, 3, 5)$, $S_3 = (1, 2, 3, 6)$, $S_4 = (1, 2, 3, 7)$, $S_5 = (1, 2, 3, 8)$, $S_6 = (1, 2, 4, 5)$, $S_7 = (1, 2, 4, 6)$, ..., $S_{15} = (1, 2, 7, 8)$, ..., $S_{68} = (4, 5, 7, 8)$, $S_{69} = (4, 6, 7, 8)$, and $S_{70} = (5, 6, 7, 8)$. In the 70 sets of $n = 4$ cultivars, each set of 4 cultivars occurs once in the design, each triplet of cultivars occurs $m - 3 = 5$ times, each pair of cultivars occurs $(m-2)(m-3)/2 = 15$ times, and each cultivar occurs $(m-1)(m-2)(m-3)/6 = 35$ times. There are $m!/3!(m-4)! = m(m-1)(m-2)(m-3)/6 = 280$ responses $= 4(70)$ responses available for this design. The total 280 degrees of freedom may be partitioned as shown in Table 15.2.

15.3.5 Minimal Designs for All Effects up to KSMA Effects

In order to obtain solutions for all effects up to KSMA, $n \geq k$ and m must be large enough so that there are sufficient observations for the total number of degrees of freedom necessary for all effects in the treatment design. General formulas for m and n are given in Tables 15.1 and 15.2. The minimal treatment design for estimating all effects up to KSMA is all possible combinations of m cultivars taken $n = k$ at a time, or $m!/k!(m-k)!$ mixtures of k cultivars. Note that smaller numbers of mixtures may be possible for some m when $n > k$.

TABLE 15.1. Number of Effects, Restrictions, and Terms of kth Effects

	Number of		
Effect	Restrictions	Terms	No. Restrictions
GMA	$\sum_i (\tau_i + \delta_i) = 0$	m	1
BSMA	$\sum_{j \neq i} \beta_{i(j)} = 0$	$m(m-1)$	m
TSMA	$\sum_{h \neq ij} \pi_{i(j,h)} = 0$	$\frac{m(m-1)(m-2)}{2}$	$m(m-1)$
QSMA	$\sum_{l \neq ijk} \gamma_{i(j,k,l)} = 0$	$\frac{m(m-1)(m-2)(m-3)}{6}$	$\frac{m(m-1)(m-2)}{2}$
QuSMA	$\sum_{p \neq ijkl} \alpha_{i(j,k,l,p)} = 0$	$\frac{m(m-1)(m-2)(m-3)(m-4)}{24}$	$\frac{m(m-1)(m-2)(m-3)}{6}$
KSMA	$\sum_{z \neq \text{rest}} X_{i(jh\cdots z)}$	$\frac{m!}{(k-1)!(m-k)!}$	$\frac{m!}{(k-1)!(m-k+1)!}$

For example, we showed that m cultivars in mixtures of $n = m - 1$ cultivars were minimal designs for GMA and BSMA effects. BIBDs balanced for pairs and for triplets may be used for GMA, BSMA, and TSMA effects. BIBDs balanced for pairs, triplets, and quartets may be used when obtaining information on GMA, BSMA, TSMA, and QSMA effects. The literature on such BIBDs is scant.

It is not necessary to use BIBDs balanced for all effects for the treatment design. Use of BIBDs ensures equal precision for all effects of a given order (GMA, e.g.). Minimal designs result in the smallest number of treatments (mixtures) for obtaining solutions for the desired mixing effects. Since the number of mixtures may be large, using nonbalanced designs may make the number of mixtures even larger and, hence, would be an inefficient design.

15.4 Solutions for Parameters

Federer and Raghavarao (1987) present solutions for the case $\mu_h = \mu$ and for $n = 3$ mixtures of m cultivars. The models of the previous section differ in specificity and content. Solutions for parameters and variances are presented below for the various cases considered.

15.4.1 *Cultivar Effects for Equations (15.1)–(15.3)*

Cultivar effect treatment designs given in Table 15.1 are given below. In this case, we consider $\mu_h = \mu$. Let $\lambda_{hh'}$ be the number of times a pair of cultivars h and h' appear together in a mixture, let s be the number of mixtures in which cultivar h appears, and let v be the number of mixtures. In order to effect solutions for these

TABLE 15.2. Partitioning $m!/k!(m-k)!$ Degrees of Freedom into Degrees of Freedom for GMA, BSMA, TSMA, QSMA, and QuSMA Effects.

Source of	Degrees of freedom (m, n)								
Source of	m	$m=4$	$m=5$	$m=6$	$m=7$	$m=8$	$m=9$	$m=10$	$m=10$
Variation	n	$n=2$	$n=3$	$n=3$	$n=3$	$n=4$	$n=4$	$n=4$	$n=5$
Correction for mean	1	1	1	1	1	1	1	1	1
GMA	$m-1$	3	4	5	6	7	8	9	9
BSMA	$m(m-1)$	8	15	24	35	48	63	80	80
TSMA	$*$	—	10	30	63	112	180	270	270
QSMA	†	—	—	—	—	112	252	480	480
QuSMA	‡	—	—	—	—	—	—	—	420
Total	$nm!/k!(m-k)!$	12	30	60	105	280	504	840	1260

$m(m-1)\cdots(m-k+2)(m-2k+2)/(k-1)!$ = degrees of freedom for a k-factor mixing ability effect (KSMA).

$*$ $m(m-1)(m-4)/2$.

† $m(m-1)(m-2)(m-6)/6$.

‡ $m(m-1)(m-2)(m-3)(m-8)/24$.

models, the constraint $\sum_{h=1}^{m}(\hat{\tau}_h + \hat{\delta}_h) = 0 = \sum_{g=1}^{r} \hat{\rho}_g$ is used. Then,

$$\hat{\mu}/n = \sum_{g=1}^{r}\sum_{h=1}^{m}\sum_{h\in S_\alpha} Y_{gh(\neq h, S_\alpha)}/rsm = \bar{y}_{\cdot\cdot(\cdot,\cdot)}, \tag{15.13}$$

where S_α, $\alpha = 1, 2, \ldots, m$, is the mixture of n of m cultivars and the summation over α is the sum over the s mixtures in which h occurs. If $\mu_h = \mu$, then

$$\frac{\widehat{\mu_h + \tau_h + \delta_h}}{n} = \sum_g \sum_\alpha Y_{gh(\neq h, S_\alpha)}/rs = \bar{y}_{\cdot h(\cdot,\cdot)}, \tag{15.14}$$

$$\widehat{\tau_h + \delta_h} = n\left(\bar{y}_{\cdot h(\cdot,\cdot)} - \bar{y}_{\cdot\cdot(\cdot,\cdot)}\right). \tag{15.15}$$

The variance of a difference of two cultivar means is

$$\text{Var}\left[\widehat{\tau_h + \delta_h} - (\widehat{\tau_{h'} + \delta_{h'}})\right] = 2\sigma_\epsilon^2/r(s - \lambda_{hh'}), \tag{15.16}$$

where $\sigma_{\epsilon h}^2 = \sigma_{\epsilon h'}^2 = \sigma_\epsilon^2$. If this condition is not appropriate, individual error variances and covariances will need to be obtained.

Example 15.1. The data from Table 13.2 are used to illustrate an analysis and minimum treatment design for obtaining solutions for sole crop treatment effects and for general mixing ability (GMA) effects. The data are given in Table 15.3 and an analysis of variance is presented in Table 15.4. Note that the biomass dry weight response for the sole crops all intercropped with barley are from 12 plants. In order to obtain estimates of GMA effects, the responses for $m = n = 6$ cultivars represented in the single mixture are required. Thus, the m sole crop treatment plus the treatment with all m crops in the mixture is a minimal treatment design for estimating sole crop effects and GMA effects. Since the responses for the cultivars

TABLE 15.3. Data for Sole Crop Response (Biomass) for $m = 6$ Cultivars and for a Mixture of All $n = 6$ Cultivars from Table 13.2, Data on a 12-Plant Basis.

	Block				
Treatment	1	2	3	Total	Mean
A = *Avena fatua*	24.04	15.87	27.84	67.75	22.583
B = *Coriander sativa*	2.21	1.58	1.12	4.91	1.637
C = *Lens esculentum*	47.98	41.76	28.89	118.63	39.543
D = *Lotus corniculatus*	4.02	1.09	3.04	8.15	2.717
E = *Medicago sativa*	13.72	4.81	8.37	26.90	8.967
F = *Matricatia*	3.34	0.49	0.95	4.78	1.593
A(BCDEF)	15.84	4.80	27.00	47.64	15.880
B(ACDEF)	2.40	0.60	4.38	7.38	2.460
C(ABDEF)	81.66	64.38	84.48	230.52	76.840
D(ABCEF)	2.10	4.26	3.60	9.96	3.320
E(ABCDF)	5.58	10.56	10.92	26.06	9.020
F(ABCDE)	1.92	0.42	3.24	5.58	1.860
Totals	204.81	150.62	203.83	559.26	15.535

in the mixture are from two plants, the responses in Table 13.2 are multiplied by 6 to bring the responses to a 12-plant measurement basis.

Since each sole crop mean is $\hat{\mu} + \hat{\tau}_i$, $i = A$, B, C, D, E, F, we take the $\sum_i \tau_i = 0$ and then $\hat{\mu} = \bar{y}_{\cdot\cdot} =$ mean of m sole crops and $\hat{\tau}_i = \bar{y}_{\cdot i} - \bar{y}_{\cdot\cdot}$. For example, $\bar{y}_{\cdot\cdot} = (67.75 + 4.91 + 118.63 + 8.15 + 26.90 + 4.78 = 231.12)/18 = 12.840$, $\hat{\tau}_A = 22.583 - 12.840 = 9.743$, $\hat{\tau}_B = -11.203$, $\hat{\tau}_C = 26.703$, $\hat{\tau}_D = -10.123$, $\hat{\tau}_E = -3.873$, and $\hat{\tau}_F = -11.247$. From Table 15.4, a standard error for $\hat{\tau}_i = \sqrt{30.3997(1/3 + 1/18) = 7(30.3997)/18} = 3.438$. A standard error of a difference between two $\hat{\tau}_i$'s is $\sqrt{2(30.3997)/(r = 3)} = 4.502$. Since the response for the cultivar in the mixture are from 2 plants and those from sole crops are from 12 plants, it might be suspected that the errors would be different, but they are within sampling variation of each other, i.e., 17.2217 and 31.7839. The blocks by sole crop versus mixture response mean square, 89.3686, is larger than the other but is associated with only two degrees of freedom. These are the reasons for using the pooled mean square, 30.3997 with 22 degrees of freedom instead of the individual variances.

Solutions for the $\hat{\delta}_i$, GMA effect for cultivar i, are obtained as the difference between the means of a cultivar in a mixture and as a sole crop, i.e., $\bar{y}_{\cdot i(\text{not } i)} - \bar{y}_{\cdot i}$. The solutions are $\hat{\delta}_A = 15.880 - 22.583 = -6.703$, $\hat{\delta}_B = 0.823$, $\hat{\delta}_C = 37.297$, $\hat{\delta}_D = 0.603$, $\hat{\delta}_E = 0.053$, and $\hat{\delta}_F = 0.267$. A standard error for $\hat{\delta}_i$ is $\sqrt{2(30.3997)/(r = 3)} = 4.502$, since $\hat{\delta}_i$ is simply the difference between two means. The mean of the $\hat{\delta}_i$ is the mean of the cultivars in the mixture and their mean as sole crops, or $18.23 - 12.84 = 5.39$, with a standard error of $\sqrt{30.3997(1/18 + 1/18)} = 1.838$. Lentils, *Lens esculentum*, had a large sole crop effect plus a large GMA effect relative to the other cultivars.

15.4.2 Cultivar and BSMA Effects for Equations (15.4) and (15.5), n = 2

For $\mu_h \neq \mu_{h'} \neq \mu$ and for $\sum_{\substack{i=1\\i\neq h}}^{m} \hat{\beta}_{h(i)} = 0$, solutions for the various cultivar effects, $\mu + \tau_h + \delta_h$, and BSMA, $\beta_{h(i)}$, effects are

$$\frac{1}{2}\left(\widehat{\mu_h + \tau_h + \delta_h}\right) = \sum_{g=1}^{r} \sum_{\substack{i=1\\i\neq h}}^{m} Y_{gh(i)}/r(m-1) = \bar{y}_{\cdot h(\cdot)}, \tag{15.17}$$

$$\text{Var}\left[\frac{\widehat{\mu_h + \tau_h + \delta_h}}{2} - \left(\frac{\widehat{\mu_{h'} + \tau_{h'} + \delta_{h'}}}{2}\right)\right] = \frac{\sigma^2_{\epsilon h} + \sigma^2_{\epsilon h'}}{r(m-1)}$$

under the assumption that $E[\epsilon_{gh(h')}\epsilon_{gh'(h)}] = 0$ for $i \neq h$. This covariance term, if present, would generally be expected to be small and, hence, can be ignored. Solutions for $\beta_{h(i)}$ effects and their variances are

$$\hat{\beta}_{h(i)} = \sum_{g=1}^{r} Y_{gh(i)}/r - \bar{y}_{\cdot h(\cdot)} = \bar{y}_{\cdot h(i)} - \bar{y}_{\cdot h(\cdot)}, \tag{15.18}$$

TABLE 15.4. Analysis of Variance for Data of Table 15.3.

Source of variation	d.f.	SS	MS
Total	36	26,104.4436	
CFM	1	8,688.1041	
Blocks $= R$	2	160.2450	80.1255
Treatment $= T$	11	16,587.3007	1,507.9364
Among sole $= S$	5	3,532.4620	706.4924
Among mixtures $= M$	5	12,793.3698	2,558.6740
Sole vs. mixtures	1	261.4689	261.4689
$R \times T$	22	668.7938	30.3997
$R \times S$	10	172.2174	17.2217
$R \times M$	10	317.8392	31.7839
$R \times S$ vs. M	2	178.7372	89.3686

$$\operatorname{Var}\left(\hat{\beta}_{h(i)}\right) = \sigma^2_{\epsilon h}(m-2)/r(m-1), \tag{15.19}$$

$$\operatorname{Var}\left(\hat{\beta}_{h(i)} - \hat{\beta}_{h(i')}\right) = 2\sigma^2_{\epsilon h}/r. \tag{15.20}$$

Example 15.2. Biomass data (dry weight) for barley as a sole crop, S, and as an intercrop with $m - 1 = 6$ other cultivars are used for this example. From these data, solutions for bi-specific mixing effects, BSMA, for barley may be obtained. The data are given in Table 15.5 and an analysis of variance is presented in Table 15.6. The solutions for the BSMA effects on a 12-plant basis are obtained as the mean of cultivar i with barley minus the mean over all mixtures, 38.631, $\bar{y}_{\cdot i} - \bar{y}_{\cdot\cdot}$, $i =$A,B,C,D,E,F. Thus, $\hat{\beta}_{S(A)} = \bar{y}_{\cdot A} - \bar{y}_{\cdot S} = 34.640 - 38.631 = -3.991$, $\hat{\beta}_{S(B)} = -1.884$, $\hat{\beta}_{S(C)} = -2.808$, $\hat{\beta}_{S(D)} = 6.009$, $\hat{\beta}_{S(E)} = 0.856$, and $\hat{\beta}_{S(F)} = 1.819$. The sum of the $\beta_{S(i)}$ adds to zero within rounding error. The standard of a difference between two $\hat{\beta}_{S(i)}$ is $\sqrt{2(70.0601)/3} = 6.834$.

TABLE 15.5. Data for Mixtures of $n = 2$ of $m = 7$ Cultivars for Barley Biomass (Dry Weight) from Table 13.2. (Data on a Six-Plant Basis.)

	Block				
System	1	2	3	Total	Mean
Barley sole crop $= S$	43.90	33.09	40.98	117.97	39.323
Barley + wild oat (A)	38.16	27.48	38.28	103.92	34.640
Barley + corlander (B)	38.35	35.66	36.23	110.24	36.747
Barley + lentils (C)	33.59	33.78	40.10	107.47	35.823
Barley + trefoil (D)	30.46	49.97	53.49	133.92	44.640
Barley + alfalfa (E)	31.39	55.69	31.38	118.46	39.487
Barley + chamomile (F)	41.19	37.07	43.09	121.35	40.450
Total	257.04	272.74	283.55	813.33	38.730

TABLE 15.6. Analysis of Variance for Data of Table 15.5.

Source of variation	d.f	SS	MS
Total	21	32,595.5247	
CFM	1	31,500.2709	
Blocks $= R$	2	50.7679	25.3840
Treatment $= T$	6	203.7645	33.9608
Sole vs. rest	1	1.2321	1.2321
BSMA	5	202.5359	40.5072
$R \times T$	12	840.7214	70.0601

The difference between the mean of the six barley mixtures, 38.361, and the barley sole crop mean, 39.323, is an estimate of the GMA effect for barley, i.e., $38.361 - 39.323 = -0.692$. This value is within sampling error. The mixture mean of a cultivar minus the barley sole crop mean is an estimate of the sum of the GMA and BSMA effects. For example, for cultivar A, $\bar{y}_{\cdot S(A)} - \bar{y}_{\cdot S} = \hat{\delta}_{S(\cdot)} + \hat{\beta}_{S(A)} = 34.640 - 39.323 = (38.631 - 39.323) + (34.640 - 38.631) = -0.692 - 3.991 = -4.683$.

There is considerable variation in treatment responses from block to block. The following is a table of residuals times 21 for the data of Table 15.5:

	Block		
Treatment	1	2	3
S	138.32	−135.79	−2.53
S+A	116.13	−155.25	39.12
S+B	75.88	−27.71	−48.17
S+C	−4.69	−47.80	52.49
S+D	−255.57	107.04	148.53
S+E	−127.82	335.38	−207.56
S+F	57.75	−75.87	18.12

The two largest residuals are 335.38 and −255.57. These two values contribute $(335.38/21)^2 + (-255.57/21)^2 = 403.1650$, which is almost half of the R × T sum of squares, 840.7214. Since all mean squares are smaller than the R × T mean square, it would appear that there are outliers in this data set. Also, the mean square of 70.0601 is much larger than the R × T mean square of 30.3997 obtained for Example 15.1.

The sum of squares for BSMA in Table 15.6 is computed as $r \sum_{i=1}^{m-1} \hat{\beta}_{S(i)}^2 = 202.5359$ for i =A,B,C,D,E,F, and $m - 1 = 6$. The remaining sums of squares are computed in the usual manner; e.g., sole vs. rest sum of squares is $(813.33 - 7(117.97))^2/3(6+6^2) = (813.33 - 117.97)^2/6(3) + 117.97^2/3 - 813.33^2/3(7) = 1.2321$.

15.4.3 Cultivar and BSMA Effects for Equations (15.4) and (15.5) for $n = m - 1$

If the BIBD for $v = b = m$, $k = r = m - 1$, and $\lambda = m - 2$ is used to obtain the m mixtures of $n = m - 1$ cultivars, it is possible to obtain solutions for cultivar means and BSMA, $\beta_{h(i)}$, effects. The $\beta_{h(i)}$ solutions are obtained from a mixture of $m - 1$ cultivars and not from $n = 2$ as given above. The BSMA effect in mixtures of $n = 2$ may differ substantially from the same BSMA effect in mixtures of $n = m - 1$. However, these designs may be useful in preliminary investigations and are parsimonious since only m mixtures rather than $m(m - 1)/2$ mixtures are used in the experiment. There are $m - 1$ responses for cultivar h in each complete block of an RCBD. The $m(m - 1)$ total degrees of freedom per complete block of an RCBD are partitioned into one for the mean, $m - 1$ for cultivars, and $m(m - 2)$ for the BSMA effects. The solutions for the parameters of the following response equation are

$$Y_{gh(\neq h, S_\alpha)} = \frac{1}{m-1}\left(\mu_h + \tau_h + \delta_h + \rho_g + 2\begin{pmatrix}\beta_{h(1)} + \cdots + \beta_{h(m)} \\ i \neq h\end{pmatrix}\right) + \epsilon_{gh(\neq h), S_\alpha}, \tag{15.21}$$

where $h \in S_\alpha$, $\alpha = 1, 2, \ldots, m$ mixtures, $i \neq h$ refers to any of the other $m - 2$ cultivars in a mixture. The multiplier $2/(m - 1)$ for the $\beta_{h(i)}$ effects is included to account for the fact that an interaction was defined in Volume I, Chapter 6, for equal amounts of material for the two entities interacting. In a mixture of $n = m - 1$ cultivars, only $2/n$ of the experimental unit is occupied by the two cultivars. The other $m - 3$ cultivars occupy the rest of a mixture. Of course, the interaction could be redefined, but the present procedures put GMA and BSMA effects on the same basis as for $n = 2$ and for sole crops. The same constraints as used previously apply here as well. The cultivar h mean is

$$\frac{\widehat{\mu_h + \tau_h + \delta_h}}{(m-1)} = \sum_g \sum_\alpha Y_{gh(\neq h), S_\alpha} / r(m-1)$$
$$= \frac{Y_{\cdot h(\cdot,\cdot)}}{r(m-1)} = \bar{y}_{\cdot h(\cdot,\cdot)}, \tag{15.22}$$

where h is in mixture S_α, $\alpha = 1, \ldots, m$. The variance of a difference of two cultivar means is, $h \neq h'$,

$$\frac{1}{(m-1)^2} \operatorname{Var}\left[\widehat{\mu_h + \tau_h + \delta_h} - \left(\widehat{\mu_{h'} + \tau_{h'} + \delta_{h'}}\right)\right]$$
$$= \frac{\sigma^2_{\epsilon h} + \sigma^2_{\epsilon h'}}{r(m-1)}. \tag{15.23}$$

The BSMA effect is obtained as

$$\frac{2\hat{\beta}_{h(i)}}{(m-1)} = \frac{1}{r(m-2)} \sum_g \sum_\alpha Y_{gh(i, S_\alpha)} - \bar{y}_{\cdot h(\cdot,\cdot)}$$

$$= \bar{y}_{\cdot h(i,\cdot)} - \bar{y}_{\cdot h(\cdot,\cdot)}, \tag{15.24}$$

where hi is a member of mixture S_α. Note that cultivars h and i occur together in $\lambda = m - 2$ mixtures in each complete block of the RCBD. The variance for $\hat{\beta}_{h(i)}$ is

$$\mathrm{Var}(\hat{\beta}_{h(i)}) = \sigma^2_{\epsilon h} \left(\frac{m-2}{8r} \left(2 + (m-3)(m-2)^2\right)\right). \tag{15.25}$$

The variance of a difference between two BSMA effects is

$$\mathrm{Var}\left(\hat{\beta}_{h(i)} - \hat{\beta}_{h(i')}\right) = (m-1)^2 \sigma^2_{\epsilon h}/2r. \tag{15.26}$$

15.4.4 Cultivar, BSMA, and TSMA Effects for Equations (15.6)–(15.8), $n = 3$

Solutions for parameters of equations (15.6) to (15.8) for $v = m(m-1)(m-2)/6$ mixtures, $S_\alpha, \alpha = 1, 2, \ldots, v$, of size $n = 3$ of m cultivars are presented below. In addition to the previous constraints, the constraints $\sum_{i=1}^m \hat{\pi}_{h(i,j)} = \sum_{j=1}^m \hat{\pi}_{h(i,j)} = 0$ for $h \neq i \neq j \neq h$ are used. The cultivar means are

$$\frac{\widehat{\mu_h + \tau_h + \delta_h}}{3} = 2 \sum_{g=1}^{r} \sum_{\alpha} \frac{Y_{gh(i,j,S_\alpha)}}{r(m-1)(m-2)}$$
$$= \bar{y}_{\cdot h(\cdot,\cdot,\cdot)}. \tag{15.27}$$

The variance of a difference between two cultivar means is

$$\frac{1}{9}\mathrm{Var}\left[\widehat{\mu_h + \tau_h + \delta_h} - \left(\widehat{\mu_{h'} + \tau_{h'} + \delta_{h'}}\right)\right] \tag{15.28}$$
$$= \frac{2(m-3)\left(\sigma^2_{\epsilon h} + \sigma^2_{\epsilon h'}\right)}{r(m-1)^2(m-2)},$$

where error term effects are uncorrelated.

$$\frac{2}{3}\hat{\beta}_{h(i)} = (m-2)\left[\sum_g \sum_\alpha \frac{Y_{gh(i,j,S_\alpha)}}{r(m-2)} - \bar{y}_{\cdot h(\cdot,\cdot)}\right]$$
$$/(m-3)$$
$$= \frac{(m-2)\left(\bar{y}_{\cdot h(i,\cdot)} - \bar{y}_{\cdot h(\cdot,\cdot)}\right)}{m-3}, \tag{15.29}$$

where $hi \in S_\alpha$ (the pair hi is a member of S_α). The variance of differences between two BSMA effects is

$$\frac{4}{9}\mathrm{V}\left[\hat{\beta}_{h(i)} - \hat{\beta}_{h(i')}, i \neq i', = \frac{(m-2)\left(\bar{y}_{\cdot h(i,\cdot)} - \bar{y}_{\cdot h(i',\cdot)}\right)}{m-3}\right]$$
$$= \frac{2\sigma^2_{\epsilon h}}{r(m-3)}. \tag{15.30}$$

The solution for a TMSA effect under the above constraints is

$$\begin{aligned}\hat{\pi}_{h(i,j)} &= \bar{y}_{\cdot h(i,j)} - \bar{y}_{\cdot h(\cdot\cdot)} - \frac{2}{3}\left(\hat{\beta}_{h(i)} + \hat{\beta}_{h(j)}\right) \\ &= \bar{y}_{\cdot h(i,j)} - \frac{(m-2)\left(\bar{y}_{\cdot h(i,\cdot)} + \bar{y}_{\cdot h(j,\cdot)}\right)}{m-3} \\ &\quad + \frac{(m-1)\bar{y}_{\cdot h(\cdot,\cdot)}}{m-3}. \end{aligned} \tag{15.31}$$

Variances of differences between two TSMAs are

$$\begin{aligned}\mathrm{V}\Big[\hat{\pi}_{h(i,j)} - \hat{\pi}_{h(i,j')} &= \bar{y}_{\cdot h(i,j)} - \bar{y}_{\cdot h(i,j')} \\ &\quad + \frac{(m-2)\left(\bar{y}_{\cdot h(j'\cdot)} - \bar{y}_{\cdot h(j,\cdot)}\right)}{m-3},\ j \neq j'\Big] \\ &= \frac{2(m-4)\sigma^2_{\epsilon h}}{r(m-3)} \end{aligned} \tag{15.32}$$

and

$$\begin{aligned}\mathrm{V}\Big[\hat{\pi}_{h(i,j)} - \hat{\pi}_{h(i',j')} &= \bar{y}_{\cdot h(i,j)} - \bar{y}_{\cdot h(i',j')} - (m-2) \\ &\quad \times \left(\bar{y}_{\cdot h(i,\cdot)} + \bar{y}_{\cdot h(j,\cdot)} - \bar{y}_{\cdot h(i',\cdot)} + \bar{y}_{\cdot h(j',\cdot)}\right) \\ &\quad /\,(m-3),\ i \neq i',\, j \neq j'\Big] \\ &= \frac{2(m-5)\sigma^2_{\epsilon h}}{r(m-3)}. \end{aligned} \tag{15.33}$$

Example 15.3. BSMA and TSMA effects are obtained from the treatment design of all combinations of $m = 6$ cultivars taken $n = 3$ at a time. For cultivar E, alfalfa, there are $(m-1)(m-2)/2 = 5(4)/2 = 10$ combinations of $m-1$ cultivars taken 2 at a time. The data in Table 15.7 are obtained from Table 13.3 for cultivar E. Tables similar to Table 15.7 may be obtained for cultivars A, B, C, D, and F if desired. Cultivar E was selected to represent the estimation of $\beta_{E(i)}$, and $\pi_{E(i,j)}$ effects for i, j =A,B,C,D,F.

From the data in Table 13.3, $Y_{\cdot E(\cdot,\cdot)} = 9.51 + 4.56 + \cdots + 8.33 = 92.34$, $\bar{y}_{\cdot E(\cdot,\cdot)} = 2Y_{\cdot E(\cdot,\cdot)}/r(m-1)(m-2) = 3.078$, $Y_{\cdot E(A,\cdot)} = 9.51 + 4.56 + 7.20 + 16.44 = 37.71$, and $\bar{y}_{\cdot E(A,\cdot)} = Y_{\cdot E(A,\cdot)}/r(m-2) = 37.71/3(4) = 3.142$. The

TABLE 15.7. Totals, Means, Effects, and Variances for the Data of Table 13.3 for Cultivar E. ($Y_{\cdot E(i,j)}$ Totals Are Given in Table 13.3.)

Totals	Means	Effects
$Y_{\cdot E(\cdot,\cdot)} = 92.34$	$\bar{y}_{\cdot E(\cdot,\cdot)} = 3.078$	$\bar{y}_{\cdot E(\cdot,\cdot)} = \hat{\mu} + \hat{\tau}_E + \hat{\delta}_E$
$Y_{\cdot E(A,\cdot)} = 37.71$	$\bar{y}_{\cdot E(A,\cdot)} = 3.142$	$\hat{\beta}_{E(A)} = 0.128$
$Y_{\cdot E(B,\cdot)} = 38.32$	$\bar{y}_{\cdot E(B,\cdot)} = 3.193$	$\hat{\beta}_{E(B)} = 0.230$
$Y_{\cdot E(C,\cdot)} = 31.24$	$\bar{y}_{\cdot E(C,\cdot)} = 2.603$	$\hat{\beta}_{E(C)} = -0.950$
$Y_{\cdot E(D,\cdot)} = 37.17$	$\bar{y}_{\cdot E(D,\cdot)} = 3.098$	$\hat{\beta}_{E(D)} = 0.040$
$Y_{\cdot E(F,\cdot)} = 40.24$	$\bar{y}_{\cdot E(F,\cdot)} = 3.353$	$\hat{\beta}_{E(F)} = 0.550$

TSMA Effects

$\hat{\pi}_{E(A,B)} = -0.147$	$\hat{\pi}_{E(A,C)} = -1.010$	$\hat{\pi}_{E(A,D)} = -0.790$	$\hat{\pi}_{E(A,F)} = 1.950$
	$\hat{\pi}_{E(B,C)} = 0.465$	$\hat{\pi}_{E(B,D)} = 0.412$	$\hat{\pi}_{E(B,F)} = -0.728$
		$\hat{\pi}_{E(C,D)} = 1.072$	$\hat{\pi}_{E(C,F)} = -0.525$
			$\hat{\pi}_{E(D,F)} = -0.695$

$\text{Var}(\hat{\beta}_{E(i)} - \hat{\beta}_{E(j)}, i \neq j) = 1.7500$
$\text{Var}(\hat{\pi}_{E(i,j)} - \hat{\pi}_{E(i,k)}, j \neq k) = 1.5556$
$\text{Var}(\hat{\pi}_{E(i,j)} - \hat{\pi}_{E(k,l)}, i, j \neq k, l) = 0.7778$
Standard Error of a difference between two $\hat{\beta}_{E(i)} = 1.32$
Standard error of a difference between $\hat{\pi}_{E(i,j)}$ and $\hat{\pi}_{E(i,j')} = 1.247$

remaining means are computed in a similar manner. $\bar{y}_{\cdot E(\cdot,\cdot)} = 3.078$ is an estimate of $\hat{\mu} + \hat{\tau}_E + \hat{\delta}_E$. If $\sum_{k=1}^{m} \hat{\tau}_i = 0$, i =A,B,C,D,E,F, then $\sum_{h=1}^{6} \bar{y}_{\cdot h(\cdot,\cdot)}/6 = \hat{\mu} + \bar{\delta}$, where $\bar{\delta} = \sum_{h=1}^{m} \hat{\delta}_h/m$ and h =A,B,C,D,E,F.

From equation (15.29), $\hat{\beta}_{E(i)} = \frac{3}{2}\left(\frac{m-2}{m-3}\right)\left(\bar{y}_{\cdot E(i,\cdot)} - \bar{y}_{\cdot E(\cdot,\cdot)}\right)$; $\hat{\beta}_{E(A)} = \frac{3}{2}\left(\frac{6-2}{6-3}\right) \cdot (3.142 - 3.078) = 0.128$. The remaining $\hat{\beta}_{E(i)}$ are obtained in a similar manner and are given in Table 15.7. The sum of the $\hat{\beta}_{E(i)}$ is zero within rounding errors. From equation (15.32), solutions for $\pi_{E(i,j)}$ are obtained; for example, $\hat{\pi}_{E(i,j)} = \bar{y}_{\cdot E(i,j)} + \frac{m-1}{m-3}\,\bar{y}_{\cdot E(\cdot,\cdot)} - \frac{m-2}{m-3}\,\left(\bar{y}_{\cdot E(i,\cdot)} + \bar{y}_{\cdot E(j,\cdot)}\right)$ and $\hat{\pi}_{E(A,B)} = 9.51 + 5(3.078)/3 - 4(3.142 + 3.193)/3 = -0.147$. The remaining $\hat{\pi}_{E(i,j)}$ are computed in a similar manner and are given in Table 15.7. The sum of the $\hat{\pi}_{E(i,j)}$ summed over $j \neq E$ or i is zero. This serves as a computational check.

An analysis of variance for the data used to obtain Table 15.7 is given in Table 15.8. The R, T, and $R \times T$ sums of squares are the usual ones for a randomized complete block design. The sum of squares for BSMA is computed as $r2^2(m-3)\sum_{i=1}^{m-1} \hat{\beta}_{E(i)}^2/3^2 = 4(0.128^2 + \cdots + 0.550^2) = 5.1035$ with $m - 2 = 4$ degrees of freedom. The TSMA sum of squares is computed as $r\sum\sum_{i<j} \hat{\pi}_{E(i,j)}^2 = 3[(-0.147)^2 + (-1.010)^2 + \cdots + (-0.695)^2] = 3(8.292096) = 24.8763$ with $(m-1)(m-4)/2 = 5$ degrees of freedom. The sum of the BSMA and TSMA sums of squares adds to the sum of squares for the $(m-1)(m-2)/2 = 10$ mixtures, sum of squares.

TABLE 15.8. Analysis of Variance for Cultivar Data from Table 13.5 with Monocrop and Six-Cultivar Mixture Responses Included.

Source of variation	d.f.	SS	MS
Total	36	438.702	
CFM	1	338.070	
Blocks $= R$	2	1.124	0.562
Treatment $= T$	11	30.011	2.728
Mixtures of three	9	29.978	3.331
BSMA	4	5.104	1.276
TSMA	5	24.876	4.975
Sole vs. three	1	0.023	0.023
Rest vs. six	1	0.011	0.011
$R \times T$	22	69.496	3.159
$R\times$ mixtures of three	18	62.978	3.500
Remainder	4	6.518	1.630

Since the analysis of variance in Table 15.8 is the same as the one in Table 13.5, the sole crop E and the mixture of all $m = 6$ cultivars were included with the 10 mixtures of 3, making 12 treatments as given in Table 12.3. The sums of squares for sole versus and for rest versus the mixture of six mixtures of three, are computed as indicated in Chapter 13. All other sums of squares are computed using standard procedures.

The variance of a difference between two $\hat{\beta}_{E(i)}$'s is obtained from equation (15.30) as $2(3/2)^2\sigma^2_{\epsilon E}/r(m-3) = 2(9/4)(3.500)/3(6-3) = 1.7500$ and a standard error of $\sqrt{1.7500} = 1.32$. The variance of a difference between $\hat{\pi}_{E(i,j)} - \hat{\pi}_{E(i,j')}$, $j = j'$, is obtained from equation (15.32) as $2(m-4)\sigma^2_{\epsilon E}/r(m-3) = 2(6-4)(3.500)/3(6-3) = 1.5556$ with a standard error of $\sqrt{1.5556} = 1.247$. The variance of a difference between two $\hat{\pi}_{E(i,j)} - \hat{\pi}_{E(k,l)}$, $i, j \neq k, l$, is obtained from equation (15.33) as $2(m-5)\sigma^2_{\epsilon E}/r(m-3) = 2(6-5)(3.500)/3(6-3) = 0.7778$ and a standard error of 0.882. $\hat{\pi}_{E(A,F)}$ differs from all other except $\hat{\pi}_{E(A,C)}$, $\hat{\pi}_{E(B,C)}$, $\hat{\pi}_{E(B,D)}$, and $\hat{\pi}_{E(C,D)}$ by more than two standard errors. $\hat{\pi}_{E(C,D)}$ differs from $\hat{\pi}_{E(B,F)}$ by more than two standard errors.

15.4.5 Cultivar, BSMA, TSMA, and QSMA Effects for Equations (15.9)–(15.12), $n = 4$

Solutions for the parameters of equations (15.9) to (15.12) for $m(m-1)(m-2)(m-3)/24$ mixtures, S_α, $\alpha = 1, 2, \ldots, v$, of size $n = 4$ are given here. In addition to the constraints used above, the following ones are added:

$$\sum_{i=1}^{m} \gamma_{h(i,j,k)} = \sum_{j=1}^{m} \gamma_{h(i,j,k)} = \sum_{k=1}^{m} \gamma_{h(i,j,k)} = 0 \tag{15.34}$$

for $h \neq i \neq j \neq k$.

The cultivar means are given in Appendix 15.1 and solutions for effects are

$$\hat{\mu} + \hat{\tau}_h + \hat{\delta}_h = 6 \sum_{g=1}^{r} \sum_{\alpha} Y_{gh(i,j,k)} / r(m-1)(m-2)(m-3)$$

$$= \bar{y}_{\cdot h(\cdot,\cdot,\cdot)}, \tag{15.35}$$

$$\frac{1}{2}\hat{\beta}_{h(i)} = \left(\frac{m-2}{m-4}\right)\left(\bar{y}_{\cdot h(i,\cdot,\cdot)} - \bar{y}_{\cdot h(\cdot,\cdot,\cdot)}\right), \tag{15.36}$$

$$\frac{3}{4}\hat{\pi}_{h(i,j)} = \left(\frac{m-3}{m-5}\right)\bar{y}_{\cdot h(i,j,\cdot)} + \left(\frac{m-1}{m-5}\right)$$

$$\times\, \bar{y}_{\cdot h(\cdot,\cdot,\cdot)} - \left(\frac{(m-2)}{(m-5)}\right)\left(\bar{y}_{\cdot h(i,\cdot,\cdot)} + \bar{y}_{\cdot h(j,\cdot,\cdot)}\right), \tag{15.37}$$

and

$$\hat{\gamma}_{h(i,j,k)} = \bar{y}_{\cdot h(i,j,k)} - \frac{(m-1)(m-2)}{(m-4)(m-5)}\bar{y}_{\cdot h(\cdot,\cdot,\cdot)}$$

$$+ \left(\frac{(m-2)(m-3)}{(m-4)(m-5)}\right)\left(\bar{y}_{\cdot h(i,\cdot,\cdot)} + \bar{y}_{\cdot h(j,\cdot,\cdot)} + \bar{y}_{\cdot h(k,\cdot,\cdot)}\right)$$

$$- \left(\frac{m-3}{m-5}\right)\left(\bar{y}_{\cdot h(i,j,\cdot)} + \bar{y}_{\cdot h(i,k,\cdot)} + \bar{y}_{\cdot h(j,k,\cdot)}\right). \tag{15.38}$$

Variances for differences of estimated effects are obtained in Appendices 15.1 and 15.2 and are not repeated here. Use may be made of a computer program in Appendix 15.2 to obtain numerical coefficients of $\sigma^2_{\epsilon h}$ for specific values of m. Extending the programs, expressions for mixtures of $n = 5, 6, \ldots$ may be obtained as described for $n = 4$.

Example 15.4. Biomass data for barley from the experiment described in Chapter 12 are presented in Table 15.9, along with the totals and means for the $(m-1)(m-2)(m-3)/6 = 20$ combinations of mixtures of $n = 4$ for $m = 7$ cultivars. The means for cultivar S = barley with cultivar A averaged over all other cultivars is

$$\bar{y}_{\cdot S(A\cdot\cdot)} = (126.24 + 120.25 + 138.85 + 124.93 + 133.09$$

$$+ 106.96 + 116.28 + 126.74 + 138.72 + 116.36)/3(10)$$

$$= 41.6140,$$

where $r = 3$ replicates and $(m-2)(m-3)/2 = 10$ combinations of $m-2$ cultivars, where S and A are both in the mixture of $n = 4$ cultivars. The means $\bar{y}_{\cdot S(B\cdot\cdot)}$, $\bar{y}_{\cdot S(C\cdot\cdot)}$, $\bar{y}_{\cdot S(D\cdot\cdot)}$, $\bar{y}_{\cdot S(E\cdot\cdot)}$, and $\bar{y}_{\cdot S(F\cdot\cdot)}$ are obtained in a similar manner. Of course, $\bar{y}_{\cdot S(\cdots)} = 2,539.01/60 = 42.3168$ is the mean over all 20 combinations. The mean $\bar{y}_{\cdot S(AB\cdot)} = (126.24 + 120.25 + 138.85 + 124.93)/3(7-3) = 42.5225$. The remaining means $\bar{y}_{\cdot S(ij\cdot)}$ are obtained similarly.

An analysis of variance for the barley biomass is presented in Table 15.10. The standard randomized complete block analysis is used to obtain the block,

TABLE 15.9. Data for Barley (Dry Weight) Responses in a Mixture of $n = 4$ of $m = 7$ Cultivars for Totals, Means, BSMA, TSMA, and QSMA Effects.

	Block			Total	Mean
Mixture	1	2	3	$Y_{\cdot S(ijk)}$	$\bar{y}_{\cdot S(ijk)}$
S + A + B + C	33.99	46.22	46.03	126.24	42.080
S + A + B + D	36.04	46.61	37.60	120.25	40.083
S + A + B + E	37.97	51.60	49.28	138.85	46.283
S + A + B + F	42.61	41.73	40.59	124.93	41.643
S + A + C + D	34.12	42.43	56.54	133.09	44.363
S + A + C + E	34.40	32.23	40.33	106.96	35.653
S + A + C + F	38.32	39.69	38.27	116.28	38.760
S + A + D + E	32.77	46.08	47.89	126.74	42.247
S + A + D + F	42.50	47.24	48.98	138.72	46.240
S + A + E + F	33.31	39.68	43.37	116.36	38.787
S + B + C + D	47.67	48.18	36.69	132.54	44.180
S + B + C + E	31.78	34.39	37.72	103.89	34.630
S + B + C + F	34.19	45.61	55.18	134.98	44.933
S + B + D + E	41.88	41.21	49.18	132.27	44.090
S + B + D + F	40.69	60.72	46.91	148.32	49.440
S + B + E + F	43.17	36.70	44.69	124.56	41.520
S + C + D + E	34.39	54.95	34.61	123.95	41.317
S + C + D + F	43.84	34.55	48.75	127.14	42.380
S + C + E + F	46.68	43.37	42.23	132.28	44.093
S + D + E + F	39.36	48.37	42.93	130.66	43.553
Totals	769.68	881.56	887.77	2539.01	42.317

$\bar{y}_{\cdot S(AB\cdot)} = 42.5225$	$\bar{y}_{\cdot S(\cdot\cdot\cdot)} = 42.3168$	$\bar{y}_{\cdot S(A\cdot\cdot)} = 41.6140$	$\bar{y}_{\cdot S(B\cdot\cdot)} = 42.8943$	$\bar{y}_{\cdot S(C\cdot\cdot)} = 41.2450$
$\bar{y}_{\cdot S(AC\cdot)} = 40.2142$	$\bar{y}_{\cdot S(BC\cdot)} = 41.4708$		$\bar{y}_{\cdot S(D\cdot\cdot)} = 43.7893$	$\bar{y}_{\cdot S(E\cdot\cdot)} = 41.2173$
$\bar{y}_{\cdot S(AD\cdot)} = 43.2333$	$\bar{y}_{\cdot S(BD\cdot)} = 44.4483$	$\bar{y}_{\cdot S(CD\cdot)} = 43.0600$		$\bar{y}_{\cdot S(F\cdot\cdot)} = 43.1410$
$\bar{y}_{\cdot S(AE\cdot)} = 40.7425$	$\bar{y}_{\cdot S(BE\cdot)} = 41.6308$	$\bar{y}_{\cdot S(CE\cdot)} = 38.9233$	$\bar{y}_{\cdot S(DE\cdot)} = 42.8017$	
$\bar{y}_{\cdot S(AF\cdot)} = 41.3575$	$\bar{y}_{\cdot S(BF\cdot)} = 44.3992$	$\bar{y}_{\cdot S(CF\cdot)} = 42.5567$	$\bar{y}_{\cdot S(DF\cdot)} = 45.4033$	$\bar{y}_{\cdot S(EF\cdot)} = 41.9883$
TSMA effects			BSMA effects	
$\hat{\pi}_{S(AB)} = 0.9662$		$\hat{\beta}_{S(A)} = -2.3437$	$\hat{\beta}_{S(B)} = 1.9250$	$\hat{\beta}_{S(C)} = -3.5727$
$\hat{\pi}_{S(AC)} = 0.3084$	$\hat{\pi}_{S(BC)} = -0.6083$		$\hat{\beta}_{S(D)} = 4.9083$	$\hat{\beta}_{S(E)} = -3.6650$
$\hat{\pi}_{S(AD)} = -0.1217$	$\hat{\pi}_{S(BD)} = -1.1493$	$\hat{\pi}_{S(CD)} = 0.6462$		$\hat{\beta}_{S(F)} = 2.7473$
$\hat{\pi}_{S(AE)} = 1.8095$	$\hat{\pi}_{S(BE)} = -0.0893$	$\hat{\pi}_{S(CE)} = -1.8117$	$\hat{\pi}_{S(DE)} = 0.0497$	
$\hat{\pi}_{S(AF)} = -2.9628$	$\hat{\pi}_{S(BF)} = 0.8807$	$\hat{\pi}_{S(CF)} = 1.4651$	$\hat{\pi}_{S(DF)} = 0.5750$	$\hat{\pi}_{S(EF)} = 0.0417$
GSMA effects				
$\hat{\gamma}_{S(ABC)} = 1.2587$	$\hat{\gamma}_{S(ACE)} = -2.1030$	$\hat{\gamma}_{S(BCD)} = 1.0665$	$\hat{\gamma}_{S(BEF)} = -1.9253$	
$\hat{\gamma}_{S(ABD)} = -4.2502$	$\hat{\gamma}_{S(ACF)} = -1.0808$	$\hat{\gamma}_{S(BCE)} = -3.1485$	$\hat{\gamma}_{S(CDE)} = 1.0014$	
$\hat{\gamma}_{S(ABE)} = 3.9930$	$\hat{\gamma}_{S(ADE)} = -0.8236$	$\hat{\gamma}_{S(BCF)} = 0.8236$	$\hat{\gamma}_{S(CDF)} = -3.9930$	
$\hat{\gamma}_{S(ABF)} = -1.0014$	$\hat{\gamma}_{S(ADF)} = 3.1488$	$\hat{\gamma}_{S(BDE)} = 1.0807$	$\hat{\gamma}_{S(CEF)} = 4.2504$	
$\hat{\gamma}_{S(ACD)} = 1.9253$	$\hat{\gamma}_{S(AEF)} = -1.0662$	$\hat{\gamma}_{S(BDF)} = 2.1031$	$\hat{\gamma}_{S(DEF)} = -1.2586$	

treatment, and $R \times T$ sums of squares. For this experiment, we use the $R \times T$ mean square as the error mean square of 36.2350 to obtain the various variances and standard errors of differences for the BSMA, TSMA, and QSMA effects. The partitioning of the treatment sum of squares into sums of squares for BSMA,

TABLE 15.10. Analysis of Variance for Data of Table 15.9

Source of variation	d.f.	SS	MS
Total	60	109,982.2607	
CFM	1	107,442.8630	
Blocks $= R$	2	441.6824	220.8412
Treatments $= T$	19	720.7838	37.9360
BSMA	5	301.6325	60.3265
TSMA	9	73.4120	8.1569
QSMA	5	345.7366	69.1473
$R \times T$	38	1,376.9315	36.2350

TSMA, and QSMA effects is described later. The degrees of freedom of these sums of squares are obtained from Table 15.2. For this example, the BSMA sum of squares has $m-2=5$ degrees of freedom for the barley biomass yields. There are $m = 7$ cultivars resulting in the $m(m-2)$ degrees of freedom from Table 15.2. Since we are using the barley cultivar only, there are $m-2=5$ degrees of freedom for the BSMA sum of squares. The TSMA sum of squares is associated with $(m-1)(m-4)/2 = 9$ degrees of freedom. The QSMA sum of squares is associated with $(m-1)(m-2)(m-6)/6 = 5$ degrees of freedom.

The BSMA effects, $\hat{\beta}_{S(i)}$, i =A,B,C,D,E,F, are computed as $\hat{\beta}_{S(i)} = 2(m-2)\left(\bar{y}_{\cdot S(i\cdot\cdot)} - \bar{y}_{\cdot S(\cdots)}\right)/(m-4)$, equation (15.35), which for $i = A$ is $\hat{\beta}_{S(A)} = 2(7-2)(41.6140 - 42.3168)/(7-4) = -2.3427$. The remaining $\hat{\beta}_{S(i)}$ are computed likewise, and their sum is zero within rounding errors. The variance of a difference between two $\hat{\beta}_{S(i)}$, equation (15.87), is 16(error mean square)/$r(m-3)(m-4) = 16(36.2350)/36 = 16.1044$, and a standard error of a difference of 4.01. We note that $\hat{\beta}_{S(D)} = 4.9083$ is approximately two standard errors larger than $\hat{\beta}_{S(C)}$ and $\hat{\beta}_{S(E)}$; the former has a positive effect and the latter two a negative effect on the biomass of barley. Cultivar D is trefoil, cultivar C is lentils, and cultivar E is alfalfa. The cultivar F, chamomile, BSMA effect is about 1.5 standard errors larger than the BSMA effects for alfalfa, E, and lentils, C. Since the biomass of chamomile is relatively small, this could account for its positive effect. The ordering of the BSMA effects here is different from Example 15.3, again pointing up the fact that generalizations for effects over all mixture sizes is not an appropriate method for interpreting intercropping data.

The TSMA effects are computed from equation (15.36) as

$$\frac{4}{3}\left[\frac{(m-3)}{(m-5)}\bar{y}_{\cdot(ij\cdot)} + \frac{(m-1)}{(m-5)}\bar{y}_{\cdot S(\cdots)} - \frac{(m-2)}{(m-5)}\left(\bar{y}_{\cdot S(i\cdot\cdot)} + \bar{y}_{\cdot S(j\cdot\cdot)}\right)\right].$$

For example, $\hat{\pi}_{S(AB)} = \frac{2}{3}[4(42.5225) + 6(42.3168) - 5(41.6140 + 42.8943)] = 0.9662$. The remaining $\hat{\pi}_{S(ij)}$ are computed in a similar manner. As a computational check, $\sum_{\substack{j=1 \\ \neq S,A}}^{m} \hat{\pi}_{S(ij)} = 0$ for each i. The variance of a difference between $\hat{\pi}_{S(ij)}$ and $\hat{\pi}_{S(ij')}$, $j \neq j'$ [equation (15.88)] is 32(error mean square)/9$r(m-3)(m-5) = 32(36.2350)/9(3)(4)(2) = 5.3681$ and a standard error of a difference of

$\sqrt{5.3681} = 2.32$. The largest differences are found for $\hat{\pi}_{S(AF)} = -2.9628$ and $\hat{\pi}_{S(AE)} = 1.8095$ and $\hat{\pi}_{S(DF)} = 1.4651$. These differences are approximately two standard errors of a difference apart. The variance of a difference between $\hat{\pi}_{S(ij)}$ and $\hat{\pi}_{S(i'j')}, i, j \neq i', j'$, is 32 (error mean square)/$9r(m-3) = 32(36.2350)/9(3)(4) = 10.7363$ and a standard error of a difference of 3.28. $\hat{\pi}_{S(AF)}$ is a little more than one standard error of a difference lower than $\hat{\pi}_{S(CD)}$. None of the contrasts approach significance at the 10% level.

From equation (15.37), we compute the QSMA effects as

$$\begin{aligned}\hat{\lambda}_{S(ijk)} = \bar{y}_{\cdot S(ijk)} &- \frac{(m-3)}{(m-5)}\left(\bar{y}_{\cdot S(ij\cdot)} + \bar{y}_{\cdot S(ik\cdot)} + \bar{y}_{\cdot S(jk\cdot)}\right)\\ &+ \frac{(m-2)(m-3)}{(m-4)(m-5)}\left(\bar{y}_{\cdot S(i\cdot\cdot)} + \bar{y}_{\cdot S(j\cdot\cdot)} + \bar{y}_{\cdot S(k\cdot\cdot)}\right)\\ &- \frac{(m-1)(m-2)}{(m-4)(m-5)}\bar{y}_{\cdot S(\cdot\cdot\cdot)},\end{aligned}$$

which, for $\hat{\lambda}_{S(ABC)}$, is equal to $42.0800 - 2(42.5225 + 40.2142 + 41.4708) + (10/3)(41.6140 + 42.8943 + 41.2450) - 5(42.3168) = 1.2587$. Using the above equation, the remaining $\hat{\lambda}_{S(ijk)}$ were obtained. Note that the $\sum^{m}_{\substack{k=1\\ \neq S,i, \text{ or } j}} \hat{\lambda}_{S(ijk)} = 0$ for each ij. This serves as a computational check. The variance of a difference between two $\hat{\lambda}_{S(ijk)}$ and $\hat{\lambda}_{S(ijk'}$, $k \neq k'$, from equation (15.90), is $2(m-6)(\text{EMS})/r(m-4) = 2(36.2350)/3(3) = 8.0522$ and standard error of a difference of $(8.0522)^{1/2} = 2.84$. Several differences are larger than two standard errors, e.g., $\hat{\lambda}_{S(ABD)}$ and $\hat{\lambda}_{S(ABE)}$, and $\hat{\lambda}_{S(ABD)}$ and $\hat{\lambda}_{S(ADF)}$. The variance of a difference between $\hat{\lambda}_{S(ijk)}$ and $\hat{\lambda}_{S(ij'k')}$, $j, k \neq j'k'$, is $2(m-6)^2(\text{EMS})/r(m-4)(m-5)$, which is $2(36.2350)/3(3)(2) = 4.0261$ and a standard error of a difference of 2.01. Several differences, e.g., $\hat{\lambda}_{S(ABE)}$ and $\hat{\lambda}_{S(ADF)}$ exceed two standard errors of a difference. The variance of a difference between $\hat{\lambda}_{S(ijk)}$ and $\hat{\lambda}_{S(i'j'k')}, i, j, k \neq i', j', k'$, is $[2(m-6)^2 + 4]\text{EMS}/r(m-4)(m-5) = 6(36.2350)/3(3)(2) = 12.0783$ and a standard error of $(12.0783)^{1/2} = 3.48$. Again, several differences, e.g., $\hat{\lambda}_{S(ABD)}$ and $\hat{\lambda}_{S(CEF)}$ and $\hat{\lambda}_{S(CDF)}$ and $\hat{\lambda}_{S(ABE)}$, exceed two standard errors of a difference.

In this experiment, significant (5% level) differences existed for the $\hat{\beta}_{S(i)}$ and $\hat{\lambda}_{S(ijk)}$ but not for the $\hat{\pi}_{S(ij)}$. An experimenter would desire large positive $\hat{\beta}_{S(i)}$, $\hat{\pi}_{S(ij)}$, and $\hat{\lambda}_{S(ijk)}$. For example, the effects of barley biomass of BSMA, TSMA, and QSMA effects in the four-cultivar combination of SBDF is

$$\begin{aligned}&\frac{1}{2}\left(\hat{\beta}_{S(B)} + \hat{\beta}_{S(D)} + \hat{\beta}_{S(F)}\right) + \frac{3}{4}\left(\hat{\pi}_{S(BD)} + \hat{\pi}_{S(BF)} + \hat{\pi}_{S(DF)}\right) + \hat{\lambda}_{S(BDF)}\\ &\quad = \frac{1}{2}(1.93 + 4.91 + 2.75) + \frac{3}{4}(-1.15 + 0.88 + 0.58) + 2.10\\ &\quad = 4.80 + 0.23 + 2.10 = 7.13.\end{aligned}$$

All effects except $\hat{\pi}_{S(BD)}$ have positive contributions. An experimenter would desire large positive contributions from these effects as well as a large contribution from $\left(\hat{\delta}_i + \hat{\delta}_j + \hat{\delta}_k\right)/4$, for $n = 4$, as this combination would produce large yields compared to the sole crop, e.g., barley.

15.5 Combined ANOVA for m Cultivars

For a combined analysis of variance of the 84+6(36)=300 biomass responses for the 7 cultivars, the responses should be for a constant number of plants or areas. If desired, the responses adjusted for a constant number of plants or areas may be used to obtain ANOVAs for each of the cultivars as was done in Chapter 13 for cultivars A, B, C, D, E, and F and in Examples 15.2 and 15.4 for barley biomass. Biomass responses for barley for the $v = 7$ cultivar mixtures are available and are considered to be included in the ANOVA for barley in Table 15.11. There are 21 barley responses from Example 15.2, 60 from Example 15.4, and 3 from the 7-cultivar mixture. (These 3 responses are 34.73, 40.03, and 39.78 for blocks one, two, and three, respectively), to make 84 responses. In order for a combined ANOVA to have meaning, all responses should be for the same characteristic biomass here. Grain yield (Table 12.2) and number of tillers (Table 15.12) responses relate to barley only and not to responses for the other cultivars.

The sums of squares for treatments for each cultivar may be partitioned into component parts as was done in the ANOVAs presented in Chapters 13 and 15. The combined ANOVA given in Table 15.11 is easily expanded to include the partitioning of the treatment sums of squares for each of the cultivars. The ANOVA in Table 15.13 is useful to summarize the data from an intercropping experiment such as the one described in Chapter 12.

From the four examples and the problems presented in the next section, it is demonstrated that some creativity will be involved to obtain a complete and comprehensive analysis for an intercropping experiment. Each experiment will present different aspects so that "one size fits all" cannot apply to analyses for this type of experiment. Sometimes, different analyses will be needed for the different characters measured from the same mixture experiment. Thus, in considering such experiments, the experimenter needs to clarify completely the goals and to be aware of anything unplanned that occurs in the experiment.

TABLE 15.11. ANOVA Degrees of Freedom for Each of the $m = 7$ Cultivars in mixtures of $n = 4$ cultivars for biomass of the mixture.

Source of		Degrees of freedom for cultivar						
Variation	Barley	A	B	C	D	E	F	Sum
Total	84	36	36	36	36	36	36	300
CFM	1	1	1	1	1	1	1	7
Block = R	2	2	2	2	2	2	2	15
Treatment = T	27	11	11	11	11	11	11	93
$R \times T$	54	22	22	22	22	22	22	186

TABLE 15.12. Number of Tillers for Barley.

	Block		
Treatment	1	2	3
Sole	20	15	17
Barley + A	17	14	19
Barley + B	17	15	18
Barley + C	15	17	18
Barley + D	15	19	21
Barley + E	17	20	15
Barley + F	17	18	18
Barley + A + B + C	15	18	18
Barley + A + B + D	16	18	15
Barley + A + B + E	17	17	18
Barley + A + B + F	17	17	18
Barley + A + C + D	18	19	18
Barley + A + C + E	18	15	18
Barley + A + C + F	17	17	16
Barley + A + D + E	14	19	22
Barley + A + D + F	18	20	17
Barley + A + E + F	16	18	17
Barley + B + C + D	22	25	18
Barley + B + C + E	18	16	18
Barley + B + C + F	19	17	20
Barley + B + D + E	17	19	20
Barley + B + D + F	18	22	19
Barley + B + E + F	22	19	19
Barley + C + D + E	18	20	15
Barley + C + D + F	18	15	16
Barley + C + E + F	20	20	17
Barley + D + E + F	17	20	20
Barley + A + B + C + D + E + F	17	21	18
Totals			

15.6 Problems

15.1 For barley grain weights from Table 12.2 for barley (S) sole, S + A, S + B, S + C, S + D, S + E, S + F, obtain estimates of BSMA, $\beta_{S(i)}$, effects as was done in Example 15.2.

15.2 For cultivars A, B, C, D, and F, compute the analyses described in Example 15.3.

15.3 For barley grain weights from Table 12.2, obtain the analysis described in Example 15.4.

15.4 Suppose than an experimenter used only the following $v = 13$ treatments: barley (S) sole, S + A, S + B, S + C, S + D, S + E, S + F, S + A + B + D, S + A + C + F, S+A+D+E, S+B+C+E, S+B+E+F, and S+C+D+F. For barley

TABLE 15.13. Combined ANOVA Degrees of Freedom for the ANOVAs in Table 15.11.

Source of variations	d.f.
Total	300
CFM within cultivar	7
CFM for ANOVA	1
Among cultivars $= C$	6
R within cultivar	14
Blocks $= R$	2
$C \times R$	12
Treatment within cultivar	93
Treatment - barley	27
Treatment - A	11
Treatment - B	11
Treatment - C	11
Treatment - D	11
Treatment - E	11
Treatment - F	11
$R \times T$ within cultivar	186
$R \times T$ - barley	54
$R \times T$ - A	22
$R \times T$ - B	22
$R \times T$ - C	22
$R \times T$ - D	22
$R \times T$ - E	22
$R \times T$ - F	22

grain weights from Table 12.2, obtain solutions for $\mu + \tau_S$, GMA effect δ_S, BSMA effects $\beta_{S(i)}$, and a residual "lack of fit" for treatments sum of squares with five degrees of freedom.

15.5 From Table 12.2 for barley grain weight, use barley sole crop (S), and the following six mixtures of $n = 4$ cultivars: S+A+B+D, S+A+C+F, S+A+D+E, S+B+C+E, S+B+E+F, and S+C+D+F, to obtain solutions for $\mu + \tau_S$, GMA effect δ_S, and BSMA effects $\beta_{S(A)}$, $\beta_{S(B)}$, $\beta_{S(C)}$, $\beta_{S(D)}$, $\beta_{S(E)}$, and $\beta_{S(F)}$.

15.6 Suppose that cultivar A, B, C, D, E, and F biomass was available only for the following 10 mixtures: ABC, ACD, ADE, AEF, ABF, BCE, BDE, BDF, CDF, and CEF. Obtain solutions for $\mu_h + \tau_h + \delta_h$ and BSMA effects $\beta_{h(i)}$ for $h, i=$ A, B, C, D, E, F. (Note: The combinations were selected so that each cultivar appears twice in a mixture with every other cultivar.)

15.7 Write a MAPLE program to obtain solutions for m cultivars in mixtures of $n = 3$ cultivars. Then, write a MATHEMATICA program to obtain variances for differences of effects as was done in Appendix 15.2.

15.8 Instead of $n = 3$, write programs for mixtures of $n = 5$ cultivars as was done for Problem 15.7.

15.9 For the data in Table 15.12 on number of barley tillers, perform the analyses described in Chapter 12, in Example 15.2, and in Example 15.4.

15.10 Using the results in Subsection 15.4.3 and the observations $Y_{gs(ABC)}$, $Y_{gs(ABD)}$, $Y_{gs(ACD)}$, and $Y_{gs(BCD)}$, the barley data from Table 12.2 obtain solutions for $\mu_s + \tau_s + \delta_s$ and $\beta_{s(i)}$ effects. Do likewise for the observations $Y_{gs(ACD)}$, $Y_{gs(ACE)}$, $Y_{gs(ADE)}$, and $Y_{gs(CDE)}$.

15.7 Literature Cited

Ezumah, H.C., W.T. Federer, and G.O. Myers (1991) Intercropping cassava with grain legumes in humid Africa. 1. cowpea yields. BU-1108-MC, Technical Report Series, Biometrics Unit, Cornell University, May.

Federer, W.T., A. Hedayat, C.C. Lowe, and D. Raghavarao (1976) Application of statistical design theory to crop estimation with special reference to legumes and mixtures of cultivars. *Agron. J.* **68**, 914-919.

Federer, W.T. (1979). Statistical design and response models for mixtures of cultivars. *Agron. J.* **71**, 701-706.

Federer, W.T. and D. Raghavarao (1987). Response models and minimal designs for mixtures of n and m items useful for intercropping and other investigations. *Biometrika* **74**, 571-577.

Hall, D.B. (1976). Mixing designs: A general model, simplifications, and some minimal designs. M.S. Thesis, Cornell University, Ithaca, NY.

Appendix 15.1

Subsection 15.4.2

The response totals for equation (15.4) are

$$Y_{\cdot h(i)} = r\left(\mu_h + \tau_h + \delta_h + \beta_{h(i)}\right)/2 + \sum_{g=1}^{r}\left(\rho_{gh} + \epsilon_{gh(i)}\right) \tag{15.39}$$

and

$$\begin{aligned} Y_{\cdot h(\cdot)} = r(m-1)\left(\mu + \tau_h + \delta_h\right)/2 + r\sum_{\substack{i=1\\ \neq h}}^{m}\beta_{h(i)} \\ + (m-1)\sum_{g=1}^{r}\rho_{gh} + \sum_{g=1}^{r}\sum_{\substack{i=1\\ \neq h}}^{m}\epsilon_{gh(i)}. \end{aligned} \tag{15.40}$$

When

$$\sum_{g=1}^{r}\rho_{gh} = \sum_{\substack{i=1\\ \neq h}}^{m}\beta_{h(i)} = E\left[\epsilon_{gh(i)}\right] = 0, \tag{15.41}$$

solutions for parameters in the above equations are

$$\frac{1}{2}\left(\hat{\mu}_h + \hat{\tau}_h + \hat{\delta}_h\right) = \sum_{g=1}^{r}\sum_{\substack{i=1\\ \neq h}}^{m} Y_{gh(i)}/r(m-1) = \bar{y}_{\cdot h(\cdot)} \tag{15.42}$$

and

$$\hat{\beta}_{h(i)} = Y_{\cdot h(i)}\big/ r - \bar{y}_{\cdot h(\cdot)} = \bar{y}_{\cdot h(i)} - \bar{y}_{\cdot h(\cdot)}. \tag{15.43}$$

When

$$E\left[\epsilon^2_{gh(i)}\right] = \sigma^2_{\epsilon h}, \tag{15.44}$$

variances for the above estimated effects and differences of effects are

$$V\left(\hat{\mu}_h + \hat{\tau}_h + \hat{\delta}_h = \bar{y}_{\cdot h(\cdot)}\right) = 4\sigma^2_{\epsilon h}\big/ r(m-1), \tag{15.45}$$

$$V\left(\hat{\beta}_{h(i)} = \bar{y}_{\cdot h(i)} - \bar{y}_{\cdot h(\cdot)}\right) = (m-2)\sigma^2_{\epsilon h}\big/ r(m-1), \tag{15.46}$$

and

$$V\left(\beta_{h(i)} - \beta_{h(i')} = \bar{y}_{\cdot h(i)} - \bar{y}_{\cdot h(i')}, i \neq i'\right) = 2\sigma^2_{\epsilon h}\big/ r. \tag{15.47}$$

When the ρ_{gh} are independently distributed with mean zero and variance $\sigma^2_{\rho h}$, (15.45) becomes

$$V\left(\hat{\mu}_h + \hat{\tau}_h + \hat{\delta}_h\right) = 4\left(\sigma^2_{\epsilon h} + \sigma^2_{\rho h}\right)\big/ r(m-1). \tag{15.48}$$

Subsection 15.4.4

For response equation (15.6) and for crop h, the totals are

$$Y_{\cdot h(i,j)} = r\left[\frac{\mu_h + \tau_h + \delta_h}{3} + \frac{2}{3}\left(\beta_{h(i)} + \beta_{h(j)}\right) + \pi.f._{\cdot h(i,j)}\right] + \sum_{g=1}^{r}(\rho_{gh} + \epsilon_{gh(i,j)}), \tag{15.49}$$

$$Y_{\cdot h(i,\cdot)} = \frac{r(m-2)}{3}\left(\mu_h + \tau_h + \delta_h + 2\beta_{h(i)}\right) + (m-2)\sum_{g=1}^{r}\rho_{gh} + r\sum_{\substack{j=1 \\ j\neq i \text{ or } h}}^{m}\left(\frac{2}{3}\beta_{h(j)} + \pi.f._{\cdot h(i,j)}\right) + \sum_{g=1}^{r}\sum_{\substack{j=1 \\ j\neq i \text{ or } h}}^{m}\epsilon_{gh(i,j)}, \tag{15.50}$$

$$Y_{\cdot h(\cdot,j)} = \frac{r(m-2)}{3}\left(\mu_h + \tau_h + \delta_h + 2\beta_{h(j)}\right) + (m-2)\sum_{g=1}^{r}\rho_{gh} + r\sum_{\substack{i=1 \\ i\neq j \text{ or } h}}^{m}\left(\frac{2}{3}\beta_{h(i)} + \pi.f._{\cdot h(i,j)}\right) + \sum_{g=1}^{r}\sum_{\substack{i=1 \\ i\neq h \text{ or } j}}^{m}\epsilon_{gh(i,j)}, \tag{15.51}$$

and

$$Y_{\cdot h(\cdot,\cdot)} = \frac{r(m-1)(m-2)}{2(3)}(\mu_h + \tau_h + \delta_h) + \frac{(m-1)(m-2)}{2}\sum_{g=1}^{r}\rho_{gh} + \frac{2r}{3}(m-2)\left(\sum_{\substack{i=1 \\ i\neq h \text{ or } j}}^{m}\beta_{h(i)} + \sum_{\substack{j=1 \\ j\neq h \text{ or } i}}^{m}\beta_{h(j)}\right) + r\sum\sum_{\substack{i<j=2 \\ h\neq i \text{ or } j}}^{m}\pi.f._{\cdot h(i,j)} + \sum_{g=1}^{r}\sum\sum_{\substack{i<j=2 \\ h\neq i \text{ or } j}}^{m}\epsilon_{gh(i,j)}. \tag{15.52}$$

When

$$\sum_{g=1}^{r}\rho_{gh} = \sum_{\substack{i=1 \\ i\neq h}}^{m}\beta_{h(i)} = \sum_{\substack{j=1 \\ j\neq h}}^{m}\beta_{h(j)} = \sum_{\substack{i=1 \\ i\neq h \text{ or } j}}^{m}\pi.f._{\cdot h(i,j)} = \sum_{\substack{j=1 \\ j\neq h \text{ or } i}}^{m}\pi.f._{\cdot h(i,j)} = 0,$$

solutions for the various quantities above are

$$\hat{\mu}_h + \hat{\tau}_h + \hat{\delta}_h = 2Y_{\cdot h(\cdot,\cdot)}\big/ r(m-1)(m-2) = \bar{y}_{\cdot h(\cdot,\cdot)}, \tag{15.53}$$

$$\hat{\beta}_{h(i)} = \frac{3}{2}\left(\frac{m-2}{m-3}\right)\left(\bar{y}_{\cdot h(i,\cdot)} - \bar{y}_{\cdot h(\cdot,\cdot)}\right), \tag{15.54}$$

$$\hat{\beta}_{h(j)} = \frac{3}{2}\left(\frac{m-2}{m-3}\right)\left(\bar{y}_{\cdot h(\cdot,j)} - \bar{y}_{\cdot h(\cdot,\cdot)}\right), \tag{15.55}$$

and

$$\pi \hat{f}_{\cdot h(i,j)} = \bar{y}_{\cdot h(i,j)} + \left(\frac{m-1}{m-3}\right)\bar{y}_{\cdot h(\cdot,\cdot)} - \left(\frac{m-2}{m-3}\right)\left(\bar{y}_{\cdot h(i,\cdot)} + \bar{y}_{\cdot h(\cdot,j)}\right). \tag{15.56}$$

Variances for the above quantities are:

$$\frac{1}{9}V\left(\hat{\mu}_h + \hat{\tau}_h + \hat{\delta}_h = \bar{y}_{\cdot h(\cdot,\cdot)}\right) = \frac{2\sigma_{\epsilon h}^2}{r(m-1)(m-2)}, \tag{15.57}$$

$$V\left[\frac{2}{3}\beta_{h(i)} = \left(\frac{m-2}{m-3}\right)\left(\bar{y}_{\cdot h(i,\cdot)} - \bar{y}_{\cdot h(\cdot,\cdot)}\right)\right]$$

$$= E\left[\left(\frac{m-2}{m-3}\right)\sum_{g=1}^{r}\sum_{\substack{j=1\\ j\neq h,i}}^{m}\epsilon_{gh(i,j)}\left(\frac{1}{r(m-2)}\right)\right.$$

$$\left. - \frac{2}{r(m-1)(m-2)}\sum_{g=1}^{r}\sum\sum_{\substack{i<j\\ h\neq i,j}}\epsilon_{gh(i,j)}\right]^2$$

$$= \frac{(m-2)\sigma_{\epsilon h}^2}{r(m-1)(m-3)}, \tag{15.58}$$

$$V\left[\frac{2}{3}\beta_{h(i)} - \frac{2}{3}\beta_{h(i')} = \left(\frac{m-2}{m-3}\right)\left(\bar{y}_{\cdot h(i,\cdot)} - \bar{y}_{\cdot h(i',\cdot)}\right), i \neq i'\right]$$

$$= E\left[\left(\frac{m-2}{m-3}\right)\left(\sum_{g=1}^{r}\sum_{\substack{j=1\\ j\neq h,i}}^{m}\epsilon_{gh(i,j)} - \sum_{g=1}^{r}\sum_{\substack{j=1\\ j\neq h,i'}}^{m}\epsilon_{gh(i',j)}\right)\frac{1}{r(m-2)}\right]^2$$

$$= \frac{2\sigma_{\epsilon h}^2}{r(m-3)}, \tag{15.59}$$

$$V\left[\pi \hat{f}_{\cdot h(i,j)} = \bar{y}_{\cdot h(i,j)} + \left(\frac{m-1}{m-3}\right)\bar{y}_{\cdot h(\cdot,\cdot)} - \left(\frac{m-2}{m-3}\right) \times \left(\bar{y}_{\cdot h(i,\cdot)} + \bar{y}_{\cdot h(\cdot,j)}\right)\right]$$

$$= E\left[\left(\sum_{g=1}^{r}\epsilon_{gh(i,j)} \Big/ r + 2\left(\frac{m-1}{m-3}\right)\sum_{g=1}^{r}\sum\sum_{\substack{i<j\\ h\neq i,j}}\epsilon_{gh(i,j)}\right)\right.$$

$$\Big/ r(m-1)(m-2)$$

$$- \left(\frac{m-2}{m-3}\right)\left(\sum_{g=1}^{r}\sum_{\substack{i=1\\ i\neq h,j}}^{m}\epsilon_{gh(i,j)} + \sum_{g=1}^{r}\sum_{\substack{j=1\\ j\neq i,h}}^{m}\epsilon_{gh(i,j)}\right)\Big/ r(m-2)\Bigg]^2$$

$$= \frac{(m-4)\sigma^2_{\epsilon h}}{r(m-3)}, \tag{15.60}$$

$$V\left(\pi\hat{.f}_{\cdot h(i,j)} - \pi\hat{.f}_{\cdot h(i,j')}\right) = V\left[\bar{y}_{\cdot h(i,j)} - \bar{y}_{\cdot h(i,j')} + \left(\frac{m-2}{m-3}\right)\left(\bar{y}_{\cdot h(\cdot,j)} - \bar{y}_{\cdot h(\cdot,j')}\right)\right]$$

$$= E\left[\sum_{g=1}^{r}\left(\epsilon_{gh(i,j)} - \epsilon_{gh(i,j')}\right)\Big/ r + \left(\frac{m-2}{m-3}\right)\right.$$

$$\times \left(\sum_{g=1}^{r}\sum_{\substack{i=1\\ i\neq h,j}}^{m}\epsilon_{gh(i,j)} - \sum_{g=1}^{r}\sum_{\substack{i=1\\ i\neq h,j'}}^{m}\epsilon_{gh(i,j')}\right)\Big/ r(m-2)\Bigg]^2$$

$$= \frac{2(m-4)\sigma^2_{\epsilon h}}{r(m-3)}, \tag{15.61}$$

and

$$V\left[\pi\hat{.f}_{\cdot h(i,j)} - \pi\hat{.f}_{\cdot h(i',j'}, i \neq i', j \neq j'\right) = V\left(\bar{y}_{\cdot h(i,j)} - \bar{y}_{\cdot h(i',j')}\right.$$

$$- \left(\frac{m-2}{m-3}\right)\left(\bar{y}_{\cdot h(i,\cdot)} + \bar{y}_{\cdot h(\cdot,j)} - \bar{y}_{\cdot h(i',\cdot)} - \bar{y}_{h(\cdot,j')}\right)\Big]$$

$$= E\left[\sum_{g}\left(\epsilon_{gh(i,j)} - \epsilon_{gh(i',j')}\right)\Big/ r - \left(\frac{m-2}{m-3}\right)\right.$$

$$\times \left(\sum_{g=1}^{r}\sum_{\substack{j=1\\ j\neq h,i}}^{m}\epsilon_{gh(i,j)} + \sum_{g=1}^{r}\sum_{\substack{i=1\\ i\neq h,j}}^{m}\epsilon_{gh(i,j)} - \sum_{g=1}^{r}\sum_{\substack{j'=1\\ j'\neq h,i'}}^{m}\epsilon_{gh(i',j')}\right.$$

$$\left. - \sum_{g=1}^{r}\sum_{\substack{i'=1\\ i'\neq h,j'}}^{m}\epsilon_{gh(i',j')}\right)\Big/ r(m-2)\Bigg]^2$$

$$= \frac{2(m-5)\sigma^2_{\epsilon h}}{r(m-3)}. \tag{15.62}$$

Subsection 15.4.5

The response totals for crop h, $h = 1, 2, \ldots, m$, using equation (15.9) for $h \neq i, j$, or k, are [let $(\mu_h + \tau_h + \delta_h)/4$ be replaced by $\mu_h + \tau_h + \delta_h$ for ease of presentation]

$$Y_{\cdot h(i,j,k)} = r\left[\mu_h + \tau_h + \delta_h + \frac{1}{2}\left(\beta_{h(i)} + \beta_{h(j)} + \beta_{h(k)}\right)\right] + \sum_{g=1}^{r} \rho_g + \frac{3r}{4}\left(\pi f_{\cdot h(i,j)} + \pi f_{\cdot h(i,k)} + \pi f_{\cdot h(j,k)}\right) + r\,\gamma_{h(i,j,k)} + \sum_{g=1}^{r} \epsilon_{gh(i,j,k)} \tag{15.63}$$

$$Y_{\cdot h(i,j,\cdot)} = r(m-3)\left[\mu_h + \tau_h + \delta_h + \frac{1}{2}\left(\beta_{h(i)} + \beta_{h(j)}\right) + \frac{3}{4}\pi f_{\cdot h(i,j)}\right] + (m-3)\sum_{g=1}^{r} \rho_g + r \sum_{\substack{k=1 \\ k \neq h,k,\text{or } j}}^{m} \left[\frac{1}{2}\beta_{h(k)} + \frac{3}{4}\left(\pi f_{\cdot h(i,k)} + \pi f_{\cdot h(j,k)}\right) + \gamma_{h(i,j,k)}\right] + \sum_{g=1}^{r} \sum_{\substack{k=1 \\ k \neq h,i,\text{or } j}}^{m} \epsilon_{gh(i,j,k)} \tag{15.64}$$

$$Y_{\cdot h(i,\cdot,k)} = r(m-3)\left[\mu_h + \tau_h + \delta_h + \frac{1}{2}\left(\beta_{h(i)} + \beta_{h(k)}\right) + \frac{3}{4}\pi f_{\cdot h(i,k)}\right] + (m-3)\sum_{g=1}^{r} \rho_g + r \sum_{\substack{j=1 \\ j \neq h,i,\text{ or } k}}^{m} \left[\frac{1}{2}\beta_{h(j)} + \frac{3}{4}\left(\pi f_{\cdot h(i,j)} + \pi f_{\cdot h(j,k)}\right) + \gamma_{h(i,j,k)}\right] + \sum_{g=1}^{r} \sum_{\substack{j=1 \\ j \neq h,i,\text{or } k}}^{m} \epsilon_{gh(i,j,k)}, \tag{15.65}$$

$$Y_{\cdot h(\cdot,j,k)} = r(m-3)\left[\mu_h + \tau_h + \delta_h + \frac{1}{2}\left(\beta_{h(j)} + \beta_{h(k)}\right) + \frac{3}{4}\pi f_{\cdot h(j,k)}\right] + (m-3)\sum_{g=1}^{r} \rho_g$$

$$+ r \sum_{\substack{i=1 \\ i \neq h, j, \text{or } k}}^{m} \left[\frac{1}{2} \beta_{h(i)} + \frac{3}{4} \left(\pi.f._{h(i,j)} + \pi.f._{h(i,k)} \right) + \gamma_{h(i,j,k)} \right]$$

$$+ \sum_{g=1}^{r} \sum_{\substack{i=1 \\ i \neq h, i, \text{ or } k}}^{m} \epsilon_{gh(i,j,k)}, \tag{15.66}$$

$$Y_{\cdot h(i,\cdot,\cdot)} = \frac{r(m-2)(m-3)}{2} \left(\mu_h + \tau_h + \delta_h + \frac{1}{2} \beta_{h(i)} \right) + \frac{(m-2)(m-3)}{2}$$

$$\times \sum_{g=1}^{r} \rho_g + \frac{r(m-2)}{2} \sum_{\substack{j=1 \\ j \neq h \text{or } i}}^{m} \beta_{h(j)} + \frac{r(m-2)}{2} \sum_{\substack{k=1 \\ k \neq h \text{or } i}}^{m} \beta_{h(k)}$$

$$+ \frac{3r(m-3)}{4} \left(\sum_{\substack{j=1 \\ j \neq h \text{ or } i}}^{m} \pi.f._{h(i,j)} + \sum_{\substack{k=1 \\ k \neq h \text{ or } i}}^{m} \pi.f._{h(i,k)} \right)$$

$$+ r \sum\sum_{\substack{j<k \\ j,k \neq h \text{or } i}} \left(\frac{3}{4} \pi.f._{h(j,k)} + \gamma_{h(i,j,k)} \right) + \sum_{g=1}^{r} \sum\sum_{\substack{j<k \\ j,k \neq h \text{or } i}} \epsilon_{gh(i,j,k)}, \tag{15.67}$$

$$Y_{\cdot h(\cdot,j,\cdot)} = \frac{r(m-2)(m-3)}{2} \left(\mu_h + \tau_h + \delta_h + \frac{1}{2} \beta_{h(j)} \right)$$

$$+ \frac{(m-2)(m-3)}{2} \sum_{g=1}^{r} \rho_g + \frac{r(m-2)}{2} \sum_{\substack{i=1 \\ i \neq h \text{or } j}}^{m} \beta_{h(i)}$$

$$+ \frac{r(m-2)}{2} \sum_{\substack{k=1 \\ k \neq h \text{or } j}}^{m} \beta_{h(k)} + \frac{3r(m-3)}{4} \sum_{\substack{i=1 \\ i \neq h \text{or } j}}^{m} \pi.f._{h(i,j)}$$

$$+ \frac{3r(m-3)}{4} \sum_{\substack{k=1 \\ k \neq h \text{or } j}}^{m} \pi.f._{h(j,k)} + r \sum\sum_{\substack{i<k \\ i,k \neq h \text{or } j}} \left(\frac{3}{4} \pi.f._{h(i,k)} + \gamma_{h(i,j,k)} \right)$$

$$+ \sum_{g=1}^{m} \sum\sum_{\substack{i<k \\ i,k \neq h \text{or } j}} \epsilon_{gh(i,j,k)}, \tag{15.68}$$

$$Y_{\cdot h(\cdot,\cdot,k)} = \frac{r(m-2)(m-3)}{2} \left(\mu_h + \tau_h + \delta_h + \frac{1}{2} \beta_{h(k)} \right) + \frac{(m-2)(m-3)}{2}$$

$$\times \sum_{g=1}^{r} \rho_g + \frac{r(m-2)}{2} \sum_{\substack{i=1 \\ i \neq h}}^{m} \beta_{h(i)} + \frac{r(m-2)}{2} \sum_{\substack{j=1 \\ j \neq h}}^{m} \beta_{h(j)}$$

$$+ \frac{3r(m-3)}{4} \sum_{\substack{i=1 \\ i \neq h \text{or } k}}^{m} \pi.f._{h(i,k)} + \frac{3r(m-3)}{4} \sum_{\substack{j=1 \\ j \neq h \text{or } k}}^{m} \pi.f._{h(j,k)}$$

$$+ r \sum_{i<j}\sum \left(\frac{3}{4} \pi.f._{h(i,j)} + \gamma_{h(i,j,k)} \right) + \sum_{g=1}^{r} \sum \sum_{\substack{i<j \\ i,j \neq h \text{or } k}} \sum \epsilon_{gh(i,j,k,)}, \quad (15.69)$$

and

$$Y_{\cdot h(\cdot,\cdot,\cdot)} = \frac{r(m-1)(m-2)(m-3)}{6} (\mu_h + \tau_h + \delta_h)$$

$$+ \frac{(m-1)(m-2)(m-3)}{6} \sum_{g=1}^{r} \rho_g$$

$$+ \frac{r(m-2)(m-3)}{4} \left[\sum_{\substack{i=1 \\ i \neq h}}^{m} \beta_{h(i)} + \sum_{\substack{j=1 \\ j \neq h}}^{m} \beta_{h(j)} + \sum_{\substack{k=1 \\ k \neq h}}^{m} \beta_{h(k)} \right]$$

$$+ \frac{3r}{4} \left[\sum \sum_{\substack{i<j \\ i,j \neq h}} \pi.f._{h(i,j)} + \sum \sum_{\substack{i<k \\ i,k \neq h}} \pi.f._{h(i,k)} + \sum \sum_{\substack{j<k \\ j,k \neq h}} \pi.f._{h(j,k)} \right]$$

$$+ r \sum \sum_{i<j<k} \sum \gamma_{h(i,j,k)} + \sum_{g=1}^{r} \sum \sum_{i<j<k} \sum \epsilon_{gh(i,j,k)}. \quad (15.70)$$

Using restrictions on the parameters from above in addition to the $\sum_{i,j,\text{or } k} \gamma_{h(i,j,k)} = 0$ for $h \neq i \neq j \neq k$ and obtaining means results in the following:

$$\bar{y}_{\cdot h(i,j,k)} = \mu_h + \tau_h + \delta_h + \frac{1}{2} \left(\beta_{h(i)} + \beta_{h(j)} + \beta_{h(k)} \right)$$

$$+ \frac{3}{4} \left(\pi.f._{h(i,j)} + \pi.f._{h(i,k)} + \pi.f._{h(j,k)} \right) + \gamma_{h(i,j,k)}, \quad (15.71)$$

$$\bar{y}_{\cdot h(i,j,\cdot)} = \mu_h + \tau_h + \delta_h + \frac{1}{2} \left(\beta_{h(i)} + \beta_{h(j)} \right) \left(\frac{m-4}{m-3} \right)$$

$$+ \frac{3}{4} \pi.f._{h(i,j)} \left(\frac{m-5}{m-3} \right), \quad (15.72)$$

$$\bar{y}_{\cdot h(i,\cdot,k)} = \mu_h + \tau_h + \delta_h + \frac{1}{2}\left(\beta_{h(i)} + \beta_{h(k)}\right)\left(\frac{m-4}{m-3}\right) + \frac{3}{4}\pi.f_{\cdot h(i,k)}\left(\frac{m-5}{m-3}\right), \tag{15.73}$$

$$\bar{y}_{\cdot h(\cdot,j,k)} = \mu_h + \tau_h + \delta_h + \frac{1}{2}\left(\beta_{h(j)} + \beta_{h(k)}\right)\left(\frac{m-4}{m-3}\right) + \frac{3}{4}\pi.f_{\cdot h(j,k)}\left(\frac{m-5}{m-3}\right), \tag{15.74}$$

$$\bar{y}_{\cdot h(i,\cdot,\cdot)} = \mu_h + \tau_h + \delta_h + \beta_{h(i)}\left(\frac{m-4}{m-2}\right), \tag{15.75}$$

$$\bar{y}_{\cdot h(\cdot,j,\cdot)} = \mu_h + \tau_h + \delta_h + \beta_{h(j)}\left(\frac{m-4}{m-2}\right), \tag{15.76}$$

$$\bar{y}_{\cdot h(\cdot,\cdot,k)} = \mu_h + \tau_h + \delta_h + \beta_{h(k)}\left(\frac{m-4}{m-2}\right), \tag{15.77}$$

and

$$\bar{y}_{\cdot h(\cdot,\cdot,\cdot)} = \mu_h + \tau_h + \delta_h. \tag{15.78}$$

Solving these equations (see Appendix 15.2), we obtain

$$\hat{\mu} + \hat{\tau}_h + \hat{\delta}_h = \bar{y}_{\cdot,\cdot,\cdot)}, \tag{15.79}$$

$$\frac{1}{2}\hat{\beta}_{h(i)} = \left(\frac{m-2}{m-4}\right)\left(\bar{y}_{\cdot h(i,\cdot,\cdot)} + \bar{y}_{\cdot h(\cdot,\cdot,\cdot)}\right), \tag{15.80}$$

$$\frac{1}{2}\hat{\beta}_{h(j)} = \left(\frac{m-2}{m-4}\right)\left(\bar{y}_{\cdot h(\cdot,j,\cdot)}\right), \tag{15.81}$$

$$\frac{1}{2}\hat{\beta}_{h(k)} = \left(\frac{m-2}{m-4}\right)\left(\bar{y}_{\cdot h(\cdot,\cdot,\cdot)}\right), \tag{15.82}$$

$$\frac{3}{4}\pi.\hat{f}_{\cdot h(i,j)} = \left(\frac{m-3}{m-5}\right)\bar{y}_{\cdot h(i,j,\cdot)} + \left(\frac{m-1}{m-5}\right)\bar{y}_{\cdot h(\cdot,\cdot,\cdot)} - \frac{(m-2)}{(m-5)}\left(\bar{y}_{\cdot h(i,\cdot,\cdot)} + \bar{y}_{\cdot h(\cdot,j,\cdot)}\right), \tag{15.83}$$

$$\frac{3}{4}\pi.\hat{f}_{\cdot h(i,k)} = \left(\frac{m-3}{m-5}\right)\bar{y}_{\cdot h(i,\cdot,k)} + \left(\frac{m-1}{m-5}\right)\bar{y}_{\cdot h(\cdot,\cdot,\cdot)} - \frac{(m-2)}{(m-5)}\left(\bar{y}_{\cdot h(i,\cdot,\cdot)} + \bar{y}_{\cdot h(\cdot,\cdot,k)}\right), \tag{15.84}$$

$$\frac{3}{4}\pi.\hat{f}_{\cdot h(j,k)} = \left(\frac{m-3}{m-5}\right)\bar{y}_{\cdot h(\cdot,j,k)}\left(\frac{m-1}{m-5}\right)\bar{y}_{\cdot h(\cdot,\cdot,\cdot)}$$

$$- \frac{(m-2)}{(m-5)} \left(\bar{y}_{\cdot h(\cdot,j,\cdot} + \bar{y}_{\cdot h(\cdot,j,\cdot)} + \bar{y}_{\cdot h(\cdot,\cdot,k)} \right), \tag{15.85}$$

and

$$\begin{aligned} \hat{\gamma}_{h(i,j,k)} = \bar{y}_{\cdot h(i,j,k)} &- \left(\frac{(m-1)(m-2)}{(m-4)(m-5)} \right) \bar{y}_{\cdot h(\cdot,\cdot,\cdot)} \\ &+ \left(\frac{(m-2)(m-3)}{(m-4)(m-5)} \right) \left(\bar{y}_{\cdot h(i,\cdot,\cdot)} + \bar{y}_{\cdot h(\cdot,j,\cdot)} + \bar{y}_{\cdot h(\cdot,\cdot,k)} \right) \\ &- \left(\frac{m-3}{m-5} \right) \left(\bar{y}_{\cdot h(i,j,\cdot)} + \bar{y}_{\cdot h(i,\cdot,k)} + \bar{y}_{\cdot h(\cdot,j,k)} \right). \end{aligned} \tag{15.86}$$

The various variances are obtained as

$$\begin{aligned} V\left(\frac{1}{2}\hat{\beta}_{h(i)} - \frac{1}{2}\hat{\beta}_{h(j)}, i \neq j \right) &= V\left[\left(\frac{m-2}{m-4} \right) \left(\bar{y}_{\cdot h(i,\cdot,\cdot)} - \bar{y}_{\cdot h(\cdot,j,\cdot)} \right) \right] \\ &= \frac{E}{r} \left[2 \sum_{g=1}^{r} \sum_{\substack{j<k \\ j,k \neq i,h}} \sum \epsilon_{gh(i,j,k)} \Big/ (m-2)(m-3) \right. \\ &\qquad \left. -2 \sum_{g=1}^{r} \sum_{\substack{i<k \\ i,k \neq j,h}} \sum \epsilon_{gh(i,j,k)} \Big/ (m-2)(m-3) \right]^2 \\ &= \times \frac{4\sigma_{\epsilon h}^2}{r(m-3)(m-4)}, \end{aligned} \tag{15.87}$$

$$\begin{aligned} V&\left(\frac{3}{4}\pi \hat{f}_{\cdot h(i,j)} - \frac{3}{4}\pi \hat{f}_{\cdot h(i,k)}, j \neq k \right) \\ &= V\left[\left(\frac{m-3}{m-5} \right) \left(\bar{y}_{\cdot h(i,j,\cdot)} - \bar{y}_{\cdot h(i,\cdot,h)} \right) \right. \\ &\qquad \left. + \frac{(m-2)}{(m-5)} \left(\bar{y}_{\cdot h(\cdot,j,\cdot)} - \bar{y}_{\cdot h(\cdot,\cdot,k)} \right) \right] \\ &= \frac{E}{r} \left[\left(\frac{m-3}{m-5} \right) \left(\sum_{g=1}^{r} \sum_{\substack{k=1 \\ k \neq j,h}}^{m} \epsilon_{gh(i,j,k)} \Big/ (m-3) \right. \right. \\ &\qquad \left. - \sum_{g=1}^{r} \sum_{\substack{j=1 \\ j \neq h,k}} \epsilon_{gh(i,j,k)} \Big/ (m-3) \right) + \left(\frac{m-2}{m-5} \right) \end{aligned}$$

$$\times \left(2 \sum_{g=1}^{r} \sum_{\substack{i<j \\ i,k \neq j,h}} \sum \epsilon_{gh(i,j,k)} \Big/ (m-2)(m-3) \right.$$

$$\left.\left. - \; 2 \sum_{g=1}^{r} \sum_{\substack{i<j \\ i,j \neq k,h}} \sum \epsilon_{gh(i,j,k)} \Big/ (m-2)(m-3) \right)\right]^2$$

$$= \frac{2\sigma_{\epsilon}^2}{r} \left[(m-4) \left\{ \left(\frac{m-3}{m-5} \right) \frac{1}{m-3} - \frac{m-2}{m-5} \right.\right.$$

$$\left.\left. \times \left(\frac{2}{(m-2)(m-3)} \right) \right\}^2 \right]$$

$$+ \left\{ \frac{(m-2)}{(m-5)} \left(\frac{2}{(m-2)(m-3)} \right) \right\}^2 \left(\frac{(m-3)(m-4)}{2} \right)$$

$$= \frac{2\sigma_{\epsilon h}^2}{r} \left(\frac{(m-4)}{(m-3)(m-5)} \right) \tag{15.88}$$

$$V\left(\frac{3}{4}\pi \hat{.f}_{\cdot h(i,j)} - \frac{3}{4}\pi \hat{.f}_{\cdot h(k,l)} i,\, j \neq k, l \right)$$

$$= V\left[\left(\frac{m-3}{m-5} \right) \left(\bar{y}_{\cdot h(i,j,\cdot)} - \bar{y}_{\cdot h(k,l)} \right) + \frac{m-2}{m-5} \right.$$

$$\left. \times \left(\bar{y}_{\cdot h(\cdot,\cdot,k)} + \bar{y}_{\cdot h(\cdot,\cdot,l)} - \bar{y}_{\cdot h(i,\cdot,\cdot)} - \bar{y}_{\cdot h(\cdot,j,\cdot)} \right) \right]$$

$$= \frac{2\sigma_{\epsilon h}^2}{r(m-3)} \tag{15.89}$$

$$V\left(\hat{\gamma}_{h(i,j,k)} - \hat{\gamma}_{h(i,j,l)},\, l \neq k \right)$$

$$= V\left[\bar{y}_{\cdot h(i,j,k)} - \bar{y}_{\cdot h(i,j,l)} - \left(\frac{m-3}{m-5} \right) \right.$$

$$\times \left(\bar{y}_{\cdot h(i,k,\cdot)} + \bar{y}_{\cdot h(j,k,\cdot)} - \bar{y}_{\cdot h(i,l,\cdot)} - \bar{y}_{\cdot h(j,l,\cdot)} \right)$$

$$\left. + \; \frac{(m-2)(m-3)}{(m-4)(m-5)} \left(\bar{y}_{\cdot h(\cdot,\cdot,k)} - \bar{y}_{\cdot h(\cdot,\cdot,l)} \right) \right]$$

$$= \frac{2(m-6)^2\sigma_{\epsilon h}^2}{r(m-4)} \tag{15.90}$$

$$V\left(\hat{\gamma}_{h(i,j,k)} - \hat{\gamma}_{h(i,l,m)} \right) = V\left[\bar{y}_{\cdot h(i,j,k)} - \bar{y}_{\cdot h(i,l,m)} \right.$$

$$- \frac{m-3}{m-5}\left(\bar{y}_{\cdot h(i,j,\cdot)} + \bar{y}_{\cdot h(i,k,\cdot)} + \bar{y}_{\cdot h(j,k,\cdot)} - \bar{y}_{\cdot h(i,l,\cdot)}\right.$$

$$\left. - \ \bar{y}_{\cdot h(i,m,\cdot)} - \bar{y}_{\cdot h(l,m,\cdot)}\right)$$

$$+ \left(\frac{(m-2)(m-3)}{(m-4)(m-5)}\right)\left(\bar{y}_{\cdot h(j,\cdot,\cdot)} + \bar{y}_{\cdot h(k,\cdot,\cdot)} - \bar{y}_{\cdot h(l,\cdot,\cdot)} - \bar{y}_{\cdot h(m,\cdot,\cdot)}\right]$$

$$= \frac{2(m-6)^2\sigma^2_{\epsilon h}}{r(m-4)(m-5)}, \tag{15.91}$$

$$V\left(\hat{\gamma}_{h(i,j,k)} - \hat{\gamma}_{h(l,m,n)}\right) = V\left[\bar{y}_{\cdot h(i,j,k)} - \bar{y}_{\cdot h(l,m,n)}\right.$$

$$- \left(\frac{m-3}{m-5}\right)\left(\bar{y}_{\cdot h(i,j,\cdot)} + \bar{y}_{\cdot h(i,k,\cdot)} + \bar{y}_{\cdot h(j,k,\cdot)} - \bar{y}_{\cdot h(l,m,\cdot)}\right.$$

$$\left. - \bar{y}_{\cdot h(l,n,\cdot)} - \bar{y}_{\cdot h(m,n,\cdot)}\right)$$

$$+ \left(\frac{(m-2)(m-3)}{(m-4)(m-5)}\right)\left(\bar{y}_{\cdot h(i,\cdot,\cdot)} + \bar{y}_{\cdot h(j,\cdot,\cdot)} + \bar{y}_{\cdot h(k,\cdot,\cdot)} - \bar{y}_{\cdot h(l,\cdot,\cdot)}\right.$$

$$\left.\left. - \ \bar{y}_{\cdot h(m,\cdot,\cdot)} - \bar{y}_{\cdot h(n,\cdot,\cdot)}\right)\right]$$

$$= \left(\frac{2(m-6)^2 + 4}{r(m-4)(m-5)}\right)\sigma^2_{\epsilon h}. \tag{15.92}$$

The algebra involved in obtaining the solutions for effects and their variances is rather tedious and error-prone. Computer programs such as MATHEMATICA and MAPLE are available to aid in performing the algebra. MAPLE and MATHEMATICA programs used to obtain the solutions in equations (15.79)–(15.86) and variances of differences of effects [equations (15.82) – (15.92)] are presented in Appendix 15.2. As n, the number of crops in a mixture, increases, the complexity of the algebra will increase. For specific values of m and n, expanded versions of the programs in Appendix 15.2 will be quite useful.

Appendix 15.2

Computer software can be of considerable help in solving for parameters of a set of equations. Such packages as MAPLE and MATHEMATICA have features which aid in the solution for parameters of a set of simultaneous equations such as equations (15.71) to (15.78). For simplicity of typing programs, let

$$u = \mu_h + \tau_h + \delta_h, \qquad a = \beta_{h(i)}/2, \qquad b = \beta_{h(j)}/2,$$

$$c = \beta_{h(k)}/2, \qquad d = 3\pi f_{\cdot h(i,j)}/4, \qquad e = 3\pi f_{\cdot h(i,k)}/4,$$

$$f = 3\pi f_{\cdot h(j,k)}/4, \qquad g = \gamma_{h(i,j,k)}$$

$$\bar{y}_{\cdot h(i,j,k)} = y1, \qquad \bar{y}_{\cdot h(i,j,\cdot)} = y2, \qquad \bar{y}_{\cdot h(i,\cdot,k)} = y3,$$

$$\bar{y}_{\cdot h(\cdot,j,k)} = y4, \qquad \bar{y}_{\cdot h(i,\cdot,\cdot)} = y5, \qquad \bar{y}_{\cdot h(\cdot,j,\cdot)} = y6.$$

$$\bar{y}_{\cdot h(\cdot,\cdot,k)} = y7, \qquad \bar{y}_{\cdot h(\cdot,\cdot,\cdot)} = y8,$$

Then, a MAPLE program for obtaining solutions for u, a, b, c, d, e, f, and g is given in Figure 15.1. (Note: The command is in bold and output is in regular type.) To obtain output, a semicolon is put at the end of a command. If no output for a command is desired, the semicolon is omitted. This is the opposite of what MATHEMATICA uses. In the h1:statement, a1 contains the parameters for which solutions are required, the y1 to y8 indicate the equations, and the word "factor" indicates that coefficients of parameters are to be factored.

In writing the MATHEMATICA program, r and h were both taken to be 1. This simplifed the summations of error terms in $S2$, $S3$, $\cdots$, $S45$, and S.

If the output of a command or statement is followed by a semicolon, the result will not be printed. Since Expand[X1^2]/.res, to Expand[X6^2]/.res outputs were desired, no semicolons were used at the end of these commands. "^2" is used to indicate the power 2 and "/.res" is used to apply the condition "res" which sets $\epsilon^2_{gh(i,j,k)} = Er$ and all cross-products of error terms with different subscripts equal to zero.

These packages may be used as an aid in obtaining variances for effects and differences between effects. The coefficients of each of the individual $\epsilon_{gh(i,j,k)}$ need to be obtained. Then, these coefficients are squared and summed. To demonstrate, consider the variances in equation (15.87) for $n = 4$ crops in a mixture of $\hat{\beta}_{h(i)} - \tilde{\beta}_{h(j)}$, $i \neq j$, which involves the expected value of $\left[\frac{m-2}{m-4}\left(\bar{y}_{\cdot h(i,\cdot,\cdot)} - \bar{y}_{\cdot h(j,\cdot,\cdot)}\right)\right]^2$, where $E\left(\epsilon_{gh(i,j,k)}\right)^2 = \sigma^2_{\epsilon h}$ and all cross-products have zero expectation; i.e., the error terms are independent. The above is equal to

$$\left[2\left(\frac{m-2}{m-4}\right)\left(\sum_{g=1}^{r}\sum_{j<k}\sum \epsilon_{gh(i,j,k)} - \sum_{r}\sum_{i<k}\sum \epsilon_{gh(i,j,k)}\right)\Big/(m-2)(m-3)\right]^2$$

$$= 4\sigma^2_{\epsilon h}\big/ r(m-3)(m-4) \qquad \text{from (15.87).}$$

```
> eqs1 :={=u + a + b + c + d + e + f + g = y1,
> u + (m - 4)*(a + b)/(m - 3) + (m - 5)*(m - 3) = y2,
> u + (m - 4)*(a + c)/(m - 3) + (m - 5)*e/(m - 3) = y3,
> u + (m - 4)*(b + c)/(m - 3) + (m - 5)*f/(m - 3) = y4,
> u + (m - 4)*a/(m - 2) = y5, u + (m - 4)*b/(m - 2) = y6,
> u + (m - 4)*c/(m - 2) = y7, u = y8}
> a1 :=solve(eqs1, u,a,b,c,d,e,f,g)
> h1 :=collect(a1,y1,y2,y3,y4,y5,y6,y7,y8,factor);
```

$$hl := \left\{ c = -\frac{(m-2)y8}{m-4} + \frac{(m-2)y7}{m-4}, \right.$$

$$d = -\frac{(m-2)y6}{m-5} - \frac{(m-2)y5}{m-5} + \frac{(m-1)y8}{m-5} - \frac{(m-3)y2}{m-5},$$

$$f = \frac{(m-3)y4}{m-5} - \frac{(m-2)y6}{m-5} + \frac{(m-1)y8}{m-5}, u = y8$$

$$a = \frac{(m-2)y5}{m-4} - \frac{(m-2)y8}{m-4}, b = \frac{(m-2)y6}{m-4} - \frac{(m-2)y8}{m-4},$$

$$e = -\frac{(m-2)y5}{m-5} + \frac{(m-3)y3}{m-5} + \frac{(m-1)y8}{m-5} - \frac{(m-2)y7}{m-5}, g = -\frac{(m-3)y4}{m-5}$$

$$+\frac{(m-2)(m-3)y6}{(m-4)(m-5)} + \frac{(m-2)(m-3)y5}{(m-4)(m-5)} - \frac{(m-3)y3}{m-5} + yl - \frac{(m-1)(m-2)y8}{(m-4)(m-5)}$$

$$\left. +\frac{(m-2)(m-3)y7}{(m-4)(m-5)} - \frac{(m-3)y2}{m-5} \right\}$$

FIGURE 15.1. MAPLE program for obtaining solution to equations (15.71)–(15.77).

The variance in (15.87) is for $(\hat{\beta}_{h(i)} - \hat{\beta}_{h(i')}/2$ but the program as written is for $\hat{\beta}_{h(i)} - \hat{\beta}_{h(i')}$ which is four times (15.87). In the MATHEMATICA program in Figure 15.2, we may find the variance for specific m for $n = 4$. The values of m used were 7, 8, 9, and 10, and from the output, the above variance was obtained for general m. For the above variance, the statement

X1=Simplify[(2*(S2-S3))/(m-3)*(m-4)]

was used to determine the coefficients for each $\epsilon_{gh(i,j,k)}$, where $S2$ and $S3$ are sums of $\epsilon_{gh(i,j,k)}$. Then, using the statement Expand[X^2]/.res, the numerical coefficient of $\sigma^2_{\epsilon h}/r$ is obtained. For $m = 7$, $h = 1$, and $n = 4$, this was

$$\frac{4(4)E_r}{(7-3)(7-4)} = \frac{4E_r}{3} = \frac{4\sigma^2_{\epsilon h}}{3r}.$$

```
m = 9;
s2=(2*Sum[e[2,j,k],{j,3,m-1},{k,j+1,m}])/((m-2)*(m-3));
s3=(2*Sum[e[2,3,k],{k,4,m}]+Sum[e[3,j,k],{j,4,m-1},
{k,j+1,m}[))/((m-2)*(m-3));
s4=(2*(e[2,3,4] + Sum[e[2,4,k],{k,5,m}]+Sum[e[3,4,k],
{k,5,m}]+Sum[e[4,j,k],{j,5,m-1},{k,j+1,m}]))/((m-2)*(m-3));
s5=(2*(e[2,3,5]+e[2,4,5]+Sum[e[4,5,k],{k,6,m}]+[3,4,5]+
Sum[e[3,5,k],{k,6,m}]+Sum[e[4,5,k],{k,6,m}]+
Sum[e[5,j,k],{j,6,m-1},{k,j+1,m}]))/((m-2)*(m-3));
s6=(2*(Sum[e[i,j,6],{i,2,4},{j,i+1,5}]+Sum[e[i,6,k],
{i,2,5},{k,7,m}]+Sum[e[6,j,k],{j,7,m-1},{k,j+1,m}]))/
((m-2)*(m-3));
s23=Sum[e[2,3,k],{k,4,m}]/(m-3);
s24=(e[2,3,4]+Sum[e[2,4,k],{k,5,m}])/(m-3);
s25=(e[2,3,5]+e[2.4.5]+Sum[e[2,5,k],{k,6,m}])/(m-3);
s26=(Sum[e[2,j,6],{j,3,5}]+Sum[e[2,6,k],{k,7,m}])/(m-3);
s34=(e[2,3,4]+Sum[e[3,4,k],{k,5,m}])/(m-3);
s35=(e[2,3,5]+e[3,4,5]+Sum[e[3,5,k],{k,6,m}])/(m-3);
s45=(e[2,4,5]+e[3,4,5]+Sum[e[4,5,k],{k,6,m}])/(m-3);
s56=(Sum[e[i,5,6],{i,2,4}]+Sum[e[5,6,k],{k,7,m}])/(m-3);
s57=(Sum[e[i,5,7],{i,2,4}]+e[5,6,7]+Sum[e[5,7,k],
{k,8,m}])/(m-3);
s67=(Sum[e[i,6,7],{i,2,5}]+Sum[e[6,7,k],{k,8,m}])/(m-3);
x1=Simplify[((m-2)*(s2-s3))/(m-4)];
x2=Simplify[((m-3)*(s23-s24))/(m-5)+((m-2)*
(s4-s3))/(m-5)];
x3=Simplify[((m-3)*(s23-s24))/(m-5)+
((m-2)*(s4+s5-s3-s2))/(m-5)];
x4=Simplify[e[2,3,4]-e[2,3,5]-((m-3)*(s24+s34-s25-
s35))/(m-5)+((m-2)*(m-3)(s4-s5))/((m-4)*(m-5))];
x5=Simplify[e[2,3,4]-e[2,5,6]-((m-3)*(s23+s24+s34-
s25-s26-s56))/(m-5)+((m-2)*(m-3)*(s2+s3+s4-s5-s6-
s7))/((m-4)*(m-5))];
res={e[i_,j_,k_] e[i_,j_,k_]->Er,
e[i_,j_,k_] e[i_,j_,h_]->0,
e[i_,j_,k_] e[i_,f_,k_]->0,
e[i_,j_,k_] e[f_,g_,h_]->0},
Expand[x1^2]/.res
Expand[x2^2]/.res
Expand[x3^2]/.res
Expand[x4^2]/.res
Expand[x6^2]/.res
```

$[281] = \frac{2\,\text{Er}}{15} \qquad [282] = \frac{5\,\text{Er}}{12} \qquad [283] = \frac{\text{Er}}{3} \qquad [284] = \frac{6\,\text{Er}}{5}$

$[285] = \frac{9\,\text{Er}}{10} \qquad [286] = \frac{11\,\text{Er}}{10}$

FIGURE 15.2. MATHEMATICA program for obtaining variances for estimated differences of parameters.

CHAPTER 16

Intercrop Mixtures When Individual Crop Responses Are Not Available

16.1 Introduction

In the previous chapter, statistical designs and analyses were presented for mixtures of n of m cultivars when the individual crop responses were available. In this chapter, statistical procedures are given for mixtures of n of m cultivars when there is a single response for the mixture. The methods presented here represent a generalization of those in Chapter 7 of Volume I. The general topic of this chapter has been considered by Federer and Raghavarao (1987), where they develop some of the required theoretical results. Their minimal designs are for m items taken $t + 1$ at a time, where t is the order of specific mixing effect being considered; a tth-order effect involves $t + 1$ of the m items. The number of cultivars must exceed $2t + 1$. They present solutions for general mixing ability effects, for bi-specific mixing ability effects, and for $t = 2$ or tri-specific mixing ability effects. Their definition of general combining ability (GMA) is denoted as cultivar effect in the following. The definition of GMA used herein removes the sole crop effect from the cultivar effect. As in the previous chapter, their notation will be followed in most cases.

Many situations dictate or are facilitated by measuring a single response for the n items in a mixture. Mixtures of wheat and other cereal cultivars, mixtures of maize varieties, mixtures of green manure crops, soil fertility measurements on the experimental unit (e.u.) where a mixture is grown, simultaneous or sequential mixing of cultivars, productivity of cropping systems, forage crop mixtures (Federer *et al.*, 1976), mixtures of drugs, cough medicines, mixtures of hospital treatments, educational systems, training systems, aerobic systems, diets, preference studies

(Raghavarao and Wiley, 1986), surveys (Federer, 1991, Chapter 5; Raghavarao and Federer, 1979; Smith *et al.*, 1975), etc. represent some of the areas wherein a single response for the mixture is available. As indicated by Balaam (1986), this situation is widespread in agriculture.

In Section 16.2, minimal treatment designs are given for estimating cultivar, general mixing ability, GMA, effects, for estimating GMA and bi-specific mixing, BSMA, effects, for estimating GMA, BSMA, and tri-specific mixing ability, TSMA, effects, for estimating GMA, BSMA, TSMA, and quat-specific mixing ability QSMA effects, and some comments are presented for the general specific mixing ability case. These effects differ from those in the previous chapter in that they are sums of the individual effects described there. In Sections 16.3–16.6, response model equations, solutions for parameters of the response models, and variances of differences of effects are developed for the four situations described in Section 16.2. Illustrative examples are also presented. In Section 16.7, the four competition models in Chapter 7 of Volume I are generalized to mixtures of n of m items. Section 16.9 contains problems for the reader's use. The algebraic developments of the statistical analyses are given, but not in the detail of Chapter 15, in Appendix 16.1. Computer programs to facilitate numerical and algebraic solutions are given with the examples and in Appendix 16.2.

16.2 Treatment Designs for Estimating up to *k*th Specific Mixing Ability Effects

The GMA, BSMA, TSMA, QSMA, etc. effects considered here are different from those in Chapter 15. There, the contribution of each cultivar in a mixture to the interaction was estimable, whereas here, as in many factorial and diallel crossing experiments, only the entire interaction is available. As in diallel crossing, the contribution of each parent to the specific combining ability interaction is not estimable. In diallel crossing experiments, general and specific combining abilities solutions are available for m equal to or greater than three. Little has been done to generalize combining ability concepts for experiments involving multiple crossing; e.g., a double-cross would be a mixture of four parents. The results given herein may be useful for this purpose. Note also that GMA may not be obtainable whereas a cultivar effect is, where this effect is equal to GMA plus the effect of the cultivar as a sole crop.

Several concepts from experiment design (ED) are useful in constructing treatment designs (TDs). Minimal numbers of mixtures required to obtain the desired solutions for various treatment designs are given in Table 16.1. Also, minimal values of m for each of the TDs are indicated. For example, if cultivar + BSMA effects are desired, m must be greater than three. If cultivar + BSMA + TSMA effects are desired, n must be greater than two and m greater than six. A partitioning of the degrees of freedom for various effects for the general case and for $m = 3$ to 11 is given in the bottom part of Table 16.1. Note the cases (—)

TABLE 16.1. Numbers of Combinations of Size m Cultivars and Degrees of Freedom for Effects.

Effect	Size of mixture	Minimal number of mixtures
Cultivar	2	$m(m$ odd$)$
	3	m
	$\vdots$	$\vdots$
	$m-1$	m
Cultivar + BSMA	$n > 1, m > 3$	$m(m-1)/2$
Cultivar + BSMA + TSMA	$n > 2, m > 5$	$m(m-1)(m-2)/6$
Cultivar + BSMA + TSMA + QSMA	$n > 3, m > 7$	$m(m-1)(m-2)(m-3)/24$

Effect	Degrees of freedom General	3	4	5	6	7	8	9	10	11
Cultivar	m-1	2	3	4	5	6	7	8	9	10
BSMA	m(m-3)/2	—	2	5	9	14	20	27	35	44
TSMA	m(m-1)(m-5)/6	—	—	—	5	14	28	48	75	110
QSMA	m(m-1)(m-2)(m-7)/24	—	—	—	—	—	14	42	90	165

where solutions for effects are not available. TDs for obtaining solutions for cultivar and GMA effects (Tables 16.2 and 16.3), for cultivar + BSMA effects (Table 16.4), for cultivar + BSMA + TSMA effects, for cultivar, BSMA, TSMA, and QSMA effects, and for additional higher-ordered mixing ability effects are discussed below.

First, minimal TDs for obtaining solutions for cultivar effects only are considered in Table 16.2. An experimenter may be interested in assessing cultivar effects for m cultivars in mixtures of size n. A situation where this could be the case would be in the early stages of a program where m is large and it is decided to study cultivar effects prior to considering interactions. If m is large, the experimenter will more than likely use $n = 2$ since v, the number of mixtures, would be equal to m, the number of cultivars. This would be similar to the top-crossing situation in plant breeding. Mixtures of sizes greater than two may also be used when $v = m$. For m cultivars, a minimal design consists of m mixtures of size $n < m$. For $n = 2$, minimal saturated designs of m mixtures are available only for odd numbers. For the TD design matrix $\mathbf{X}$, which is m by m, the matrix $\mathbf{X}'\mathbf{X}$ is a circulant matrix and is singular for m even (Searle, 1979). TDs for $m = 5$, 7, and 9 are presented in Table 16.2. For $n = 3$, minimal TDs are available for all m. The construction of these designs consists of finding a balanced incomplete block design (BIBD) or a partially balanced incomplete block design (PBIBD) which is as nearly pairwise balanced (Hedayat and Federer, 1974) as possible. Pairwise balance means that all possible ordered pairs occur an equal number, λ, of times. A cyclic latin square design is constructed and rows 1 and 2 make up the TD for $n = 2$. For $n = 3$, rows

TABLE 16.2. Minimal Cultivar Effect Treatment Designs for $n = 2, 3, 4$ and $m = 4, 5, 6, 7, 8, 9$. Columns Represent the Mixture.

$n = 2$ (designs not connected for m even):

$m = 5$	$m = 7$	$m = 9$
1 2 3 4 5	1 2 3 4 5 6 7	1 2 3 4 5 6 7 8 9
2 3 4 5 1	2 3 4 5 6 7 1	2 3 4 5 6 7 8 9 1

$n = 3$:

$m = 4$ (bib)	$m = 5$	$m = 6$	$m = 7$ (bib)
1 2 3 4	1 2 3 4 5	1 2 3 4 5 6	1 2 3 4 5 6 7
2 3 4 1	2 3 4 5 1	2 3 4 5 6 1	2 3 4 5 6 7 1
4 1 2 3	4 5 1 2 3	4 5 6 1 2 3	4 5 6 7 1 2 3
$m = 8$	$m = 9$		
1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8 9		
2 3 4 5 6 7 8 1	2 3 4 5 6 7 8 9 1		
4 5 6 7 8 1 2 3	4 5 6 7 8 9 1 2 3		

$n = 4$:

$m = 4$	$m = 5$ (bib)	$m = 6$	$m = 7$ (bib)
1 2 3 4	1 2 3 4 5	1 2 3 4 5 6	1 2 3 4 5 6 7
2 3 4 1	2 3 4 5 1	2 3 4 5 6 1	2 3 4 5 6 7 1
4 1 2 3	4 5 1 2 3	4 5 6 1 2 3	4 5 6 7 1 2 3
3 4 1 2	5 1 2 3 4	3 4 5 6 1 2	7 1 2 3 4 5 6
$m = 8$	$m = 9$		
1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8 9		
2 3 4 5 6 7 8 1	2 3 4 5 6 7 8 9 1		
4 5 6 7 8 1 2 3	4 5 6 7 8 9 1 2 3		
5 6 7 8 1 2 3 4	5 6 7 8 9 1 2 3 4		

1, 2, and 4 of the latin square are selected for the TD. For $n = 4$, an additional row of the latin square is selected in such a manner as to maintain as near pairwise balance as possible, i.e., to obtain either a BIBD or a two-associate class PBIBD. If mixtures of size $n > 4$ are desired, one may proceed in the above manner and it is always possible to obtain a two-associate class PBIBD or a BIBD by selecting appropriate rows of the cyclic latin square. For $n = m - 1$, a BIBD (Youden design) always occurs. The situations where BIBDs or Youden designs occur are indicated in Table 16.2.

Such designs as the above have been used in forage crop experiments and in survey designs to obtain answers to incriminating or embarrassing questions. Federer *et al.* (1976) and Raghavarao and Federer (1979) showed that a BIBD for m entries in incomplete blocks of size $k = n$ could be used to obtain solutions for effects such as cultivar effects. The latter reference also used supplemented block designs (SBDs) for surveys. Federer (1991), Chapter 5, demonstrated that solutions for cultivar effect parameters were possible using a PBIBD where the number of incomplete blocks, mixtures, is greater than or equal to m for $n < m$. The subject matter context in which the designs are used is irrelevant, as it is only the types of effects, cultivar here, that require solutions.

Treatment designs may be required wherein one or more standard cultivars are to be included in every mixture of size n with m other cultivars. The class of supplemented block designs, SBDs, may be used to obtain TDs for this situation (Raghavarao and Federer, 1979). A SBD takes a PBIBD or a BIBD and supplements the incomplete blocks with the standard cultivar (cultivars) and one or more blocks with all m cultivars plus the standard cultivar. Suppose cultivar A is to be included in every mixture and that $m = 3$ (cultivars 1, 2, and 3). A BIBD for $v = 3$ and $k = 2$ has the incomplete blocks (1, 2), (1, 3), and (2, 3). A SBD would contain the four mixtures (A, 1, 2), (A, 1, 3), (A, 2, 3), and (A, 1, 2, 3). SBDs for $m = 4$, 5, 6, and 7 are given for $n = 2$, 3, and 4 in Table 16.3. Raghavarao and Federer (1979) discuss SBDs in the context of survey design where one question is asked of everyone, but the remaining m questions are not. The variance for the standard cultivar cultivar effect which is included in every mixture will have a different variance than the other m cultivars. The more times a cultivar is included in a mixture, the larger will be its variance, owing to the manner of defining an effect in mixtures of size n. The opposite is true for treatments in incomplete block designs.

If it is desired to obtain solutions for cultivar and BSMA effects and ignore all other interactions as might be the case in screening cultivars to include in a mixture, the designs in Table 16.4 are, in general, minimal and optimal. They are optimal in the sense that equal pairwise balance for the cultivars is achieved in most cases. The TD for $m = 8$ is not minimal. The TD for $m = 8$ has 32 mixtures, but only 28 are required in order to obtain solutions for cultivar and BSMA effects. An unsolved problem in TD construction is to obtain optimal minimal designs allowing estimation of cultivar and BSMA effects for all m. For estimating both cultivar and BSMA effects, we note from Table 16.1 that m must be greater than three and that n must be at least two. Mixtures of $v = m(m-1)/2$ are required and all possible combinations of m items taken two at a time result in a BIBD with parameters $v = m$, $k = 2$, $b = m(m-1)/2$, $r = m - 1$, and $\lambda = r(k-1)/(v-1)$, where λ is the number of times each ordered pair of cultivars occurs in the BIBD. Use of sets of orthogonal squares is made to construct TDs for m a prime number or power of a prime number. To obtain the designs, $(m-1)/2$ orthogonal squares are required to obtain the $m(m-1)/2$ mixtures needed. Then, the first n rows of these squares provide the mixture combinations. For prime numbers, orthogonal squares are formed by writing down a cyclic latin square and then taking main right diagonals of a square to form the next orthogonal square. This method of diagonalization is the one used by Khare and Federer (1981) and Federer (1995) to construct optimal incomplete block designs for $v = sk$, i.e., s incomplete blocks of size k in each complete block. Sets of orthogonal latin squares may be found in Fisher and Yates (1938) and methods of construction have been described by Hedayat and Federer (1970, 1984), Raghavarao (1971), Raghavarao *et al.* (1986), and Denes and Keedwell (1974). To illustrate for $m = 5$ and $n = 3$, the first three rows of two orthogonal

TABLE 16.3. Supplemented Block Designs for $m = 3, 4, 5, 6$ and 7 (Columns Represent the Mixture. Common Cultivar is a.)

$\underline{n = 2}$

$m = 3$

a	a	a	a
1	2	3	1
2	3	1	2
			3

$m = 5$

a	a	a	a	a	a
1	2	3	4	5	1
2	3	4	5	1	2
					3
					4
					5

$m = 7$

a	a	a	a	a	a	a	a
1	2	3	4	5	6	7	1
2	3	4	5	6	7	1	2
							3
							4
							5
							6
							7

$m = 9$

a	a	a	a	a	a	a	a	a	a
1	2	3	4	5	6	7	8	9	1
2	3	4	5	6	7	8	9	1	2
									3
									4
									5
									6
									7
									8
									9

$\underline{n = 3}$

$m = 4$

a	a	a	a	a
1	1	1	2	1
2	2	3	3	2
3	4	4	4	3
				4

$m = 5$

a	a	a	a	a	a
1	1	1	2	2	1
2	3	3	3	4	2
4	4	5	5	5	3
					4
					5

$m = 6$

a	a	a	a	a	a	a
1	2	3	4	5	6	1
2	3	4	5	6	1	2
4	5	6	1	2	3	3
						4
						5
						6

$m = 7$

a	a	a	a	a	a	a	a
1	2	3	4	5	6	7	1
2	3	4	5	6	7	1	2
4	5	6	7	1	2	3	3
							4
							5
							6
							7

$\underline{n = 4}$

$m = 5$

a	a	a	a	a	a
1	2	3	4	5	1
2	3	4	5	1	2
3	4	5	1	2	3
4	5	1	2	3	4
					5

$m = 6$

a	a	a	a	a	a	a
1	2	3	4	5	6	1
2	3	4	5	6	1	2
3	4	5	6	1	2	3
4	5	6	1	2	3	4
						5
						6

$m = 7$

a	a	a	a	a	a	a	a
1	2	3	4	5	6	7	1
2	3	4	5	6	7	1	2
4	5	6	7	1	2	3	3
7	1	2	3	4	5	6	4
							5
							6
							7

latin squares are

1	2	3	4	5
2	3	4	5	1
3	4	5	1	2

1	2	3	4	5
3	4	5	1	2
5	1	2	3	4

This TD is a BIBD with parameters $v = 5$, $k = 3$, $b = 10$, $r = 6$, and $\lambda = 3$. Since $5(5-1)/2 = 10$, the above TD has the minimum number required and is connected for cultivar + BSMA effects and is included in Table 16.4.

The TDs for $m = 5$ were obtained as the first three rows of two orthogonal latin squares for $n = 3$ and the first four rows for $n = 4$. The TD for $m = 6$ and

TABLE 16.4. Minimal Treatment Designs for Cultivar and BSMA Effects for $n = 2, 3, 4$ and $m = 5, 6, 7, 8, 9$.

$\underline{n = 2}$ [all possible combinations of m taken two at a time, or $m(m-1)/2$ mixtures]:

$m = 5$ (BIBD $\lambda = 1$)
1 1 1 1 2 2 2 3 3 4
2 3 4 5 3 4 5 4 5 5

$m = 6$ (BIBD $\lambda = 1$)
1 1 1 1 1 2 2 2 2 3 3 3 4 4 5
2 3 4 5 6 3 4 5 6 4 5 6 5 6 6

$m = 7$ (BIBD $\lambda = 1$)
1 1 1 1 1 1 2 2 2 2 2 3 3 3 3 4 4 4 5 5 6
2 3 4 5 6 7 3 4 5 6 7 4 5 6 7 5 6 7 6 7 7

$m = 8$ (BIBD $\lambda = 1$)
1 1 1 1 1 1 1 2 2 2 2 2 2 3 3 3 3 3 4 4 4 4 5 5 5 6 6 7
2 3 4 5 6 7 8 3 4 5 6 7 8 4 5 6 7 8 5 6 7 8 6 7 8 7 8 8

$m = 9$ (BIBD $\lambda = 1$)
1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 3 3 3 3 3 3 4 4 4 4 4 5 5 5 5 6 6 6 7 7 8
2 3 4 5 6 7 8 9 3 4 5 6 7 8 9 4 5 6 7 8 9 5 6 7 8 9 6 7 8 9 7 8 9 8 9 9

$\underline{n = 3}$

$m = 5$ (BIBD $\lambda = 3$)
1 2 3 4 5 1 2 3 4 5
2 3 4 5 1 3 4 5 1 2
3 4 5 1 2 5 1 2 3 4

$m = 6$
1 1 1 1 1 1 1 2 2 2 2 2 3 3 4
2 2 3 3 3 4 5 3 3 4 4 5 4 4 5
4 5 4 5 6 6 6 5 6 5 6 6 5 6 6

$m = 7$ (BIBD $\lambda = 3$)
1 2 3 4 5 6 7 1 2 3 4 5 6 7 1 2 3 4 5 6 7
2 3 4 5 6 7 1 3 4 5 6 7 1 2 4 5 6 7 1 2 3
3 4 5 6 7 1 2 5 6 7 1 2 3 4 7 1 2 3 4 5 6

$m = 8$
1 2 3 4 5 6 7 8 1 2 3 4 5 6 7 8 1 2 3 4 5 6 7 8 1 2 3 4 5 6 7 8
2 3 4 5 6 7 8 1 2 3 4 5 6 7 8 1 3 4 5 6 7 8 1 2 3 4 5 6 7 8 1 2
4 5 6 7 8 1 2 3 5 6 7 8 1 2 3 4 8 1 2 3 4 5 6 7 7 8 1 2 3 4 5 6

$m = 9$ (BIBD $\lambda = 3$)
1 2 3 4 5 6 7 8 9 1 2 3 4 5 6 7 8 9 1 2 3 4 5 6 7 8 9 1 2 3 4 5 6 7 8 9
2 3 4 5 6 7 8 9 1 7 8 9 1 2 3 4 5 6 9 7 8 3 1 2 6 4 5 8 9 7 2 3 1 5 6 4
3 4 5 6 7 8 9 1 2 4 5 6 7 8 9 1 2 3 5 6 7 8 9 1 2 3 4 6 4 5 9 7 8 3 1 2

$\underline{n = 4}$

$m = 5$ (BIBD $\lambda = 6$)
1 2 3 4 5 1 2 3 4 5
2 3 4 5 1 3 4 5 1 2
3 4 5 1 2 5 1 2 3 4
4 5 1 2 3 2 3 4 5 1

$m = 6$ (BIBD $\lambda = 6$)
1 1 1 1 1 1 1 1 1 1 2 2 2 2 3
2 2 2 2 2 2 3 3 3 4 3 3 3 4 4
3 3 3 4 4 5 4 4 5 5 4 4 5 5 5
4 5 6 5 6 6 5 6 6 6 5 6 6 6 6

$m = 7$ (BIBD $\lambda = 6$)
1 2 3 4 5 6 7 1 2 3 4 5 6 7 1 2 3 4 5 6 7
2 3 4 5 6 7 1 3 4 5 6 7 1 2 4 5 6 7 1 2 3
3 4 5 6 7 1 2 5 6 7 1 2 3 4 7 1 2 3 4 5 6
4 5 6 7 1 2 3 7 1 2 3 4 5 6 3 4 5 6 7 1 2

$m = 8$ (not minimal)
1 2 3 4 5 6 7 8 1 2 3 4 5 6 7 8 1 2 3 4 5 6 7 8 1 2 3 4 5 6 7 8
2 3 4 5 6 7 8 1 2 3 4 5 6 7 8 1 3 4 5 6 7 8 1 2 3 4 5 6 7 8 1 2
3 4 5 6 7 8 1 2 6 7 8 1 2 3 4 5 5 6 7 8 1 2 3 4 7 8 1 2 3 4 5 6
4 5 6 7 8 1 2 3 7 8 1 2 3 4 5 6 6 7 8 1 2 3 4 5 8 1 2 3 4 5 6 7

$m = 9$
1 2 3 4 5 6 7 8 9 1 2 3 4 5 6 7 8 9 1 2 3 4 5 6 7 8 9 1 2 3 4 5 6 7 8 9
2 3 4 5 6 7 8 9 1 7 8 9 1 2 3 4 5 6 9 7 8 3 1 2 6 4 5 8 9 7 2 3 1 5 6 4
3 1 2 6 4 5 9 7 8 4 5 6 7 8 9 1 2 3 5 6 4 8 9 7 2 3 1 6 4 5 9 7 8 3 1 2
4 5 6 7 8 9 1 2 3 2 3 1 5 6 4 8 9 7 6 4 5 9 7 8 3 1 2 9 7 8 3 1 2 6 4 5

$n = 3$ is all possible combinations of 6 items taken 3 at a time, resulting in 20 mixtures, whereas only 15 are required, leaving 5 degrees of freedom for lack of fit (higher-order effects). By deleting the five combinations 123, 126, 145, 234, and 356, a saturated minimal TD is obtained as given in Table 16.4. This TD is the one illustrated in Example 16.3. For $m = 6$ and $n = 4$, all possible combinations of 6 items taken 4 at a time results in 15 mixtures, the required number to estimate cultivar and BSMA effects. For $m = 7$ and $n = 3$ and 4, the first three and four rows of three orthogonal latin squares, respectively, were used to obtain the $v = 21$ mixtures necessary to estimate cultivar and BSMA effects. For $m = 8$ and $n = 3$, the four sets of rows, (1, 2, 4), (1, 2, 5), (1, 3, 8), and (1, 3, 7) from a cyclic 8-by-8 latin square were used. For $m = 8$ and $n = 4$, the four sets of rows, (1, 2, 3, 4), (1, 2, 6, 7), (1, 3, 5, 6), and (1, 3, 7, 8), of a cyclic 8-by-8 latin square were used. The preceding two designs are not minimal. A trial and error method of deleting four combinations to obtain a connected design could be attempted. The designs were obtained by trial and error, attempting to obtain as near pairwise balance as possible. It may be possible to obtain a more balanced TD for $n = 4$. For $m = 9$ and $n = 3$ and 4, four sets of orthogonal latin squares were obtained and the first three rows were used for $n = 3$ and the first four rows for $n = 4$. For $n > 4$, a TD may be formed by taking the first n rows of $(m - 1)/2$ orthogonal latin squares for prime numbers and powers of odd prime numbers. This method may be used to obtain another design for $m = 8$ in Table 16.4.

To obtain TDs for cultivar + BSMA + TSMA effects, $m(m - 1)(m - 2)/6$ combinations are required (Table 16.1) and m must be greater than five. For $n = 3$, a BIBD with parameters $v = m, k = 3, r = (m-1)(m-2)/2, b = m(m-1)(m-2)/6$, and $\lambda = m - 2$ may be used. This BIBD is all possible combinations of m items taken three at a time. For cultivar + BSMA + TSMA + QSMA effects, $m(m - 1)(m - 2)(m - 3)/24$ combinations are required, $n = 4$, and m must be greater than seven. This TD is all possible combinations of m items taken four at a time. Minimal TDs for higher-order effects may be obtained in a similar manner. For larger values of n, TDs cannot be constructed as described for Table 16.4, as $(m - 1)(m - 2)/6$ orthogonal latin squares would be required for cultivar, BSMA, and TSMA effects and this exceeds the number possible. Some other method of construction is needed to construct minimal designs.

16.3 Response Model Equations, Solutions, and Analyses for Cultivar Effects

Response model equations and solutions for effects are treated below. In this section, TDs for cultivar effects only are treated. Second, TDs for cultivar and BSMA effects are considered in the next section. Third, TDs for cultivar, BSMA, and TSMA effects are presented in Section 16.5. In Section 16.6, TDs for cultivar, BSMA, TSMA, and QSMA effects are discussed. The notation used and results obtained by Hall (1976) and Federer and Raghavarao (1987) are used here where applicable. A RCBD is assumed throughout as the extension to other EDs is straightforward. Let the n cultivar mixture be denoted by $S_\alpha = hij\ldots$, $\alpha = 1, 2, \ldots, v$ mixtures, where $hij_{\ldots}$ denotes which cultivars are in the mixture; let

$$\text{the replicate mean be } \bar{y}_{g\cdot} = \sum_{\alpha=1}^{v} Y_{gS_a}/v,$$

$$\text{the overall mean be } \bar{y}_{\cdot\cdot} = \sum_{g=1}^{r}\sum_{\alpha=1}^{v} Y_{gS_a}/rv,$$

$$\text{the mean of the mixture } S_\alpha \text{ be } \bar{y}_{\cdot S_a} = \sum_{\alpha=1}^{v} Y_{gS_a}/r,$$

$$\text{the mean of cultivar } h \text{ be } \bar{y}_{\cdot h} = \sum_{g=1}^{r}\sum_{\alpha=1, h\epsilon S_a}^{v} Y_{gS_a}/rs,$$

h occurs in s mixtures,

the mean of the mixtures where h and i both occur be

$$\bar{y}_{\cdot hi\cdot} = \sum_{g=1}^{r}\sum_{h,i\epsilon S_a} Y_{ghij}/rp,$$

where p is the number of times the pair hi occurs in the v mixtures S_α. This notation extends directly for the triplet hij and the quartet $hijk$.

For the treatment designs in Table 16.2 from an experiment designed as a RCBD, the following response model equations for $n = 2$, $n = 3$, and $n = 4$ are used:

$$Y_{ghi} = \mu + \rho_g + (\gamma_h + \gamma_i)/2 + \epsilon_{ghi}, \tag{16.1}$$

$$Y_{ghij} = \mu + \rho_g + (\gamma_h + \gamma_i + \gamma_j)/3 + \epsilon_{ghij}, \tag{16.2}$$

$$Y_{ghijk} = \mu + \rho_g + (\gamma_h + \gamma_i + \gamma_j + \gamma_k)/4 + \epsilon_{ghijk}, \tag{16.3}$$

where μ is the overall mean effect, ρ_g is the gth replicate effect, γ_h, γ_i, γ_j, and γ_k are cultivar effects for cultivars h, i, j, and k, respectively, and the ϵ's are random error effects distributed with mean zero and variance σ_ϵ^2. The coefficients of the γ's, $1/n$, are used to put the cultivar effects on the same basis as the sole effect experimental unit. The cultivar effect $\gamma_h = \tau_h + \delta_h =$ sole crop effect plus GMA effect. The individual components of the cultivar effect cannot be obtained unless the sole crops are added to the TD. Extension to mixtures of size $n > 4$ is straightforward. Using the restrictions that the sum of the replicate effects and the sum of the cultivar effects are equal to zero, a solution for the replicate effect is the replicate mean minus the overall mean, or

$$\rho_g = \bar{y}_{g\cdot} - \bar{y}_{\ldots} \tag{16.4}$$

If the TD is a (m, n, λ) BIBD, a solution for the γ_h has a simple form; i.e., for mixtures of size n,

$$\gamma_h = n \sum_{g=1}^{r} \sum_{h \epsilon S_\alpha} \frac{Y_{ghif}/r(n-\lambda) - n\,\bar{y}_{....}}{(n-\lambda)} \tag{16.5}$$

$$= n(\bar{y}_{\cdot h\cdot\cdot} - \bar{y}_{....})/(n-\lambda),$$

where the $v = m$ mixtures are numbered from S_1 to S_v, $h \epsilon S_\alpha$ means all α for which h appears in the mixture; i.e., h is a member of S_α. The variance of a difference between two estimated cultivar means or two effects is

$$\left(\frac{n}{n-\lambda}\right)^2 \text{var}(\bar{y}_{\cdot h\cdot\cdot} - \bar{y}_{\cdot h'\cdot\cdot}) = \frac{2}{r(n-\lambda)}\sigma_\epsilon^2.$$

If the TD is a two-associate class PBIBD for the associations $\lambda_1 < \lambda_2$, a solution for the cultivar effect may be obtained using any of a number of software packages, e.g., GAUSS, MAPLE, MATHEMATICA, etc. The first package gives numeric results, and algebraic results may be obtained with the other two. Since numeric results are desired, the computations are demonstrated in the following example using GAUSS. Variances of differences between two cultivar means or effects are demonstrated in Examples 16.1, 16.2, and 16.3.

16.3.1 Example 16.1. Saturated Main Effect Treatment Design for Estimating Cultivar Effects

To illustrate the calculations for an experiment using a saturated main effect treatment (cultivar) design with $m = v$, data for biomass weight for mixtures of $n = 3$ cultivars are used. The data in Table 16.5 are from the experiment described in Chapter 12. The treatment design is from Table 16.2. In addition to the responses for the $m = 6$ mixtures, biomass weights for each of the mixtures alone with barley and the mixture with all seven cultivars are also presented. Hereafter in this example, the cultivar with barley will be referred to as "sole" crop. The treatment design for $m = 6$ cultivars in mixtures of $n = 3$ is a partially balanced incomplete block design (PBIBD) with two associate classes; i.e., a pair of cultivars either occur together $\lambda_1 = 1$ or $\lambda_2 = 2$ times in a mixture. From these data, it is possible to estimate a cultivar mean, $\mu + \tau_h + \delta_h$, and cultivar effect, $\gamma_h = \tau_h + \delta_h$, a general mean effect, μ, and GMA effect, δ_h, for these $m = 6$ cultivars. $\delta_.$ is the average of the δ_h.

First, consider an analysis for the $m = 6$ cultivars in $v = 6$ mixtures of $n = 3$. The treatment design is ABD, ACF, ADE, BCE, BEF, and CDF. A appears twice with D and once with each of the other cultivars. This association scheme holds for all cultivars. The letters refer to cultivars as follows:

A	=	wild oat	D	=	birdsfoot trefoil
B	=	coriander	E	=	alfalfa
C	=	lentils	F	=	chamomile

TABLE 16.5. Plant Biomass Responses for $m = 6$ Cultivars as "Sole" and in Mixtures of $n = 3$.

Mixture	Rep 1	Rep 2	Rep 3	Total	Mean
A	24.04	15.87	27.84	67.25	22.5833
B	2.21	1.58	1.12	4.91	1.6367
C	47.98	41.76	28.89	118.63	39.5433
D	4.02	1.09	3.04	8.15	2.7167
E	13.72	4.81	8.32	26.85	8.9500
F	3.34	0.49	0.95	4.78	1.5933
ABD	8.37	10.58	13.02	31.97	10.6567
ACF	33.47	17.85	26.74	78.06	26.0200
ADE	11.28	8.32	10.43	30.03	10.0100
BCE	26.78	28.92	25.24	80.94	26.9800
BEF	2.37	6.05	3.50	11.92	3.9733
CDF	15.89	35.98	20.02	71.89	23.9633
All six	18.25	14.17	22.27	54.69	18.2300
Total	211.72	187.47	191.38	590.57	15.1428
Total "sole"	95.31	65.60	70.16	231.07	12.8372
Total $n = 3$	98.16	107.70	98.95	304.81	16.9339
$\bar{y}_{\cdot A \cdot\cdot}$ 15.56	$\bar{y}_{\cdot B \cdot\cdot}$ 13.87	$\bar{y}_{\cdot C \cdot\cdot}$ 25.65	$\bar{y}_{\cdot D \cdot\cdot}$ 14.88	$\bar{y}_{\cdot E \cdot\cdot}$ 13.65	$\bar{y}_{\cdot F \cdot\cdot}$ 17.99

ANOVA Mixtures of $n = 3$

Source of variation	d.f.	SS	MS
Total	18	7,003.5651	
Correction for mean	1	5,161.6187	
Replicate $= R$	2	9.3443	4.6722
Mixture $= M$	5	1,464.6438	292.9288
$R \times M$	10	367.9583	36.7958

TABLE 16.6. ANOVA for the Data in Table 16.5 for "Soles" and Mixtures of $n = 3$.

Source of variation	d.f.	SS	MS
Total	36	13,817.1394	
Correction for mean	1	7,976.8715	
Replicate $= R$	2	28.2721	14.1360
Treatment $= T$	11	5,148.5379	468.0489
Among "sole" $= S$	5	3,532.8500	706.5700
Among mixture $= M$	5	1,464.6438	292.9288
S versus $M = C$	1	151.0441	151.0441
$R \times T$	22	663.4579	30.1572
$R \times S$	10	229.0941	22.9094
$R \times M$	10	367.9583	36.7958
$R \times C$	2	66.4056	33.2028

The replicate and mixture totals are given in Table 16.5. From the responses and totals, an analysis of variance (ANOVA) table is obtained and given in the Table 16.6. From the mixture means, a mean of the mixtures where cultivar A occurred is obtained as

$$\bar{y}_{\cdot A\cdot\cdot} = (31.97 + 78.06 + 30.03)/(nr = 9) = 15.5622.$$

A solution for a cultivar mean, $\mu + \tau_h + \delta_h$, and cultivar effects, $\tau_h + \delta_h - \delta.$, using the restriction that the sum of the solutions is zero, i.e., $\sum_{h=1}^{6}(\hat{\tau}_h + \hat{\delta}_h - \hat{\delta}.) = 0$, is (the means are calculated to four decimals for all examples in order to limit rounding errors):

$$\begin{aligned}
\hat{\mu} + \hat{\tau}_A + \hat{\delta}_A &= 6\bar{y}_{\cdot A\cdot\cdot} - 3\bar{y}_{\cdot D\cdot\cdot} - 2\bar{y}_{\cdot\cdot\cdot\cdot} \\
&= 6(15.5622) - 3(14.8767) - 3(16.9339) = 14.8753, \\
\hat{\mu} + \hat{\tau}_B + \hat{\delta}_B &= 6\bar{y}_{\cdot B\cdot\cdot} - 3\bar{y}_{\cdot E\cdot\cdot} - 2\bar{y}_{\cdot\cdot\cdot\cdot} = 8.3890, \\
\hat{\mu} + \hat{\tau}_C + \hat{\delta}_C &= 6\bar{y}_{\cdot C\cdot\cdot} - 3\bar{y}_{\cdot F\cdot\cdot} - 2\bar{y}_{\cdot\cdot\cdot\cdot} = 66.1018, \\
\hat{\mu} + \hat{\tau}_D + \hat{\delta}_D &= 6\bar{y}_{\cdot D\cdot\cdot} - 3\bar{y}_{\cdot A\cdot\cdot} - 2\bar{y}_{\cdot\cdot\cdot\cdot} = 8.7059, \\
\hat{\mu} + \hat{\tau}_E + \hat{\delta}_E &= 6\bar{y}_{\cdot E\cdot\cdot} - 3\bar{y}_{\cdot B\cdot\cdot} - 2\bar{y}_{\cdot\cdot\cdot\cdot} = 6.4486, \\
\hat{\mu} + \hat{\tau}_F + \hat{\delta}_F &= 6\bar{y}_{\cdot F\cdot\cdot} - 3\bar{y}_{\cdot C\cdot\cdot} - 2\bar{y}_{\cdot\cdot\cdot\cdot} = -2.9174.
\end{aligned}$$

The mean of the above six means is $\hat{\mu} + \hat{\delta}. = 16.9339$, which is also the mean of all 18 responses owing to orthogonality. The above means minus 16.9339 result in the cultivar effects:

$$\begin{aligned}
\hat{\tau}_A + \hat{\delta}_A - \hat{\delta}. &= 14.8753 - 16.9339 = -2.0586, \\
\hat{\tau}_B + \hat{\delta}_B - \hat{\delta}. &= -8.5449, \\
\hat{\tau}_C + \hat{\delta}_C - \hat{\delta}. &= +49.1679, \\
\hat{\tau}_D + \hat{\delta}_D - \hat{\delta}. &= -8.2281, \\
\hat{\tau}_E + \hat{\delta}_E - \hat{\delta}. &= -10.4853, \\
\hat{\tau}_F + \hat{\delta}_F - \hat{\delta}. &= -19.8513.
\end{aligned}$$

These effects sum to zero. It appears that there are large differences among the cultivar effects, which is substantiated by the relatively large F-value for mixtures in the ANOVA, i.e., 292.9288/36.7958 = 7.96. To obtain a variance–covariance matrix for the cultivar means, we may utilize a software package like GAUSS. A GAUSS program for performing these calculations is

```
Let X[6,6] = 1 1 0 1 0 0   1 0 1 0 0 1   1 0 0 1 1 0   0 1 1 0 1 0   0 1 0 0 1 1   0 0
1 1 0 1;
Let Y[6,1] = 31.97 78.06 30.03 80.94 11.92 71.89;
b = inv(X′ ∗ X)∗X′Y; b;
var = inv(X′ ∗ X); var;
```

X is a six-by-six zero–one matrix where a one indicates in which mixture a cultivar appears for the order of cultivars A, B, C, D, E, and F. Y is a column vector of n times a mixture total/r. n/r for this example is 3/3 = 1, resulting in the mixture totals in Y. The solutions for the cultivar means is given by b and the output from this program is the set of cultivar means given above. The variance of the entries in Y is $n^2\ \sigma_\epsilon^2/r = 3^2(36.7958)/3 = 110.3874$. This value times var from the GAUSS program results in the following variance–covariance matrix for the cultivar means or effects:

$$\begin{pmatrix} +.6296 & -.0370 & -.0370 & -.3704 & -.0370 & -.0370 \\ -.0370 & +.6296 & -.0370 & -.0370 & -.3704 & -.0370 \\ -.0370 & -.0370 & +.6296 & -.0370 & -.0370 & -.3704 \\ -.3704 & -.0370 & -.0370 & +.6296 & -.0370 & -.0370 \\ -.0370 & -.3704 & -.0370 & -.0370 & +.6296 & -.0370 \\ -.0370 & -.0370 & -.3704 & -.0370 & -.0370 & +.6296 \end{pmatrix} \times 110.3874.$$

The variance of a difference of two cultivar means which appear together in a mixture once is $110.3874\,[0.6296 + 0.6296 - (-0.0370 - 0.0370)] = 147.1685$, or a standard error of a difference of 12.13. The variance of a difference between two cultivar means which occur together twice in a mixture is $110.3874\,[0.6296 + 0.6296 - (-0.3704 - 0.3704)] = 220.7748$, or a standard error of a difference of 14.86. A 95% least significant difference is a t value for 10 degrees of freedom times the standard error of a difference, i.e., $2.23(12.13) = 27.05$ or $2.23(14.86) = 33.14$; a 95% studentized range is computed as $4.91\sqrt{147.1685/2} = 42.12$, or as $4.91\sqrt{220.7748/2} = 51.59$ (see Federer, 1967, Snedecor and Cochran, 1980, e.g.). Thus, it is seen that cultivar C is different from the rest, and the rest do not differ significantly among themselves.

The mean of the mixture with six cultivars is $\hat{\mu} + \hat{\delta}. = 18.2300$, which is close to the estimate from the 18 responses mean of 16.9339. From the responses for the "sole" crops, an estimate of the mean effect μ is 12.8372. An estimate of $\delta.$ is $16.9339 - 12.8372 = 4.0967$. An ANOVA for the 36 responses for "sole" and mixtures of 3 is given in Table 16.6. Here, we see that the mixture of three means is significantly different from the mean of the "sole" crops, i.e., 151.0441/30.1572=5.01 and $F_{.05}(1, 30) = 4.17$. There are large differences among the "sole" and among the mixtures of three as indicated by the large F-values obtained from Table 16.6.

Inclusion of "sole" crops allows estimation of "sole" crops effects τ_h and GMA effect δ_h. The "sole" crop effects from the means in Table 16.5 are

$$\hat{\tau}_A = 22.5833 - 12.8372 = 9.7461,$$

$$\hat{\tau}_B = 1.6367 - 12.8372 = -11.2005,$$

$$\hat{\tau}_C = 39.5433 - 12.8372 = 26.7061,$$

$$\hat{\tau}_D = 2.7167 - 12.8372 = -10.1205,$$

$$\hat{\tau}_E = 8.9500 - 12.8372 = -3.8872,$$

$$\hat{\tau}_F = 1.5933 - 12.8372 = -11.2439.$$

The GMA effects δ_h are

$$\hat{\delta}_A = (\hat{\tau}_A + \hat{\delta}_A - \hat{\delta}.) + \hat{\delta}. - \hat{\tau}_A = -2.0583 + 4.0967 - 9.7461 = -7.7077,$$

$$\hat{\delta}_B = (\hat{\tau}_B + \hat{\delta}_B - \hat{\delta}.) + \hat{\delta}. - \hat{\tau}_B = -8.5450 + 4.0967 + 11.2005 = 6.7522,$$

$$\hat{\delta}_C = (\hat{\tau}_C + \hat{\delta}_C - \hat{\delta}.) + \hat{\delta}. - \hat{\tau}_C = 49.1683 + 4.0967 - 26.7061 = 26.5589,$$

$$\hat{\delta}_D = (\hat{\tau}_D + \hat{\delta}_D - \hat{\delta}.) + \hat{\delta}. - \hat{\tau}_D = -8.2283 + 4.0967 + 10.1205 = 5.9889,$$

$$\hat{\delta}_E = (\hat{\tau}_E + \hat{\delta}_E - \hat{\delta}.) + \hat{\delta}. - \hat{\tau}_E = -10.4850 + 4.0967 + 3.8872 = -2.5011,$$

$$\hat{\delta}_F = (\hat{\tau}_F + \hat{\delta}_F - \hat{\delta}.) + \hat{\delta}. - \hat{\tau}_F = -19.8517 + 4.0967 + 11.2439 = -4.5111$$

A variance of the difference between two "sole" crop τ_h effects is 2(22.9094)/($r = 3$) $= 15.2729$. This variance added to either of the two differences between cultivar effects is the variance of a difference between two δ_h. The square root of the resulting sum is the standard error of a difference between two GMA effects.

16.3.2 Example 16.2. Supplemented Block Design for Estimating Cultivar Effects

The computations for obtaining an ANOVA, cultivar means, and cultivar effects is illustrated using the data in Table 16.7 for $m = 7$ cultivars in mixtures of $n = 3$ and $n = 7$ for a supplemented block treatment design. The design is obtained from Table 16.3 for the mixture size $n = 3$. The data represent total biomass for the mixture. An ANOVA for these data is the standard one for RCBD and is presented at the bottom of Table 16.7. Here, we note that there are relatively large differences among the mixture means, $F = 181.6249/27.2871 = 6.66$ as compared to $F_{.01}(6, 12) = 4.82$. A 95% lsd (least significant difference) is computed as $t_{.05}$(12 d.f.)(standard error of a difference) = $2.18\sqrt{2(27.2871)/3} = 9.31$, and a studentized range (hsd) is computed as $q_{.05(12,7)}\sqrt{27.2871/3} = 4.95(3.016) = 14.95$. (See, e.g., Federer, 1967, Snedecor and Cochran, 1980.) Differences among the three highest mixture means for ACFG, BCEG, and CDFG do not exceed the lsd. Mixture CDFG has the highest mean and it is more than one lsd above all seven, BEFG, ADEG, and ABDG means. The mixture mean for ACFG is more than one lsd higher than the means of BEFG, ADEG, and ABDG. BCEG mean is more than one lsd above the means of ABDG and BEFG. The seven-mixture mean is more than one lsd above the BEFG mean. Comparisons with an hsd may also be made, but either procedure indicates differences.

Instead of concentrating on the mixture means as above, a study of cultivar means, $\mu + \tau_h + \delta_h$, and cultivar effects, $\tau_h + \delta_h$, may be desired. Cultivar effects on biomass production of a mixture are often enlightening. Solutions for cultivar means are obtained by a procedure similar to that used for Example

TABLE 16.7. Supplemented Block Treatment Design for $m = 7$ Cultivars for Biomass.

Mixture	Replicate 1	Replicate 2	Replicate 3	Total	Mean
ABDG	44.41	57.19	50.62	152.22	50.7400
ACFG	71.79	57.54	65.01	194.34	64.7800
ADEG	44.05	54.40	58.32	156.77	52.2567
BCEG	58.56	63.31	62.96	184.83	61.6100
BEFG	45.54	42.75	48.19	136.48	45.4933
CDFG	59.73	70.53	68.77	199.03	66.3433
All 7	52.98	54.20	62.05	169.23	56.4100
Total	377.06	399.92	415.92	1,192.90	56.8048

ANOVA

Source of variation	d.f.	SS	MS
Total	21	69,288.5504	
Correction for mean	1	67,762.4005	
Replicate $= R$	2	108.9847	54.4924
Mixture $= M$	6	1,089.7495	181.6249
$R \times M$	12	327.4457	27.2871

16.1. A MAPLE program for obtaining algebraic solutions to the cultivar means is:

eqs1: $\{(a+b+d+g)/4 = y_{.abdg}, (a+c+f+g)/4 = y_{.acfg}, (a+d+e+g)/4 = y_{.adeg}, (b+c+e+g)/4 = y_{.bceg}, (b+e+f+g)/4 = y_{.befg}, (c+d+f+g)/4 = y_{.cdfg}, (a+b+c+d+e+f+g)/7 = y_{.abcdefg}$;
a1:solve(eqs1, $\{a, b, c, d, e, f, g\}$)

In the above $h = a, \cdots, g$ for the cultivar mean $\mu + \tau_h + \delta_h = h$. The solutions obtained arranged in matrix form are

$$\begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 4 & 16 & 4 & -8 & -8 & -20 & 21 \\ 16 & -8 & -20 & 4 & 4 & -8 & 21 \\ -8 & 4 & -8 & 16 & -20 & 4 & 21 \\ 4 & -20 & 4 & -8 & -8 & 16 & 21 \\ -20 & -8 & 16 & 4 & 4 & -9 & 21 \\ -8 & 4 & -8 & -20 & 16 & 4 & 21 \\ 12 & 12 & 12 & 12 & 12 & 12 & -63 \end{pmatrix} \times \begin{pmatrix} y_{.abdg} \\ y_{.acfg} \\ y_{.adeg} \\ y_{.bceg} \\ y_{.befg} \\ y_{.cdfg} \\ y_{.abcdefg} \end{pmatrix}$$

An easy algebraic check is to obtain the solution for cultivar G mean, g, and this coincides with the above solution.

To obtain the variance covariance matrix for the above solutions, a GAUSS program as follows was used:

```
format 2,4;
let X[7,7] = .25 .25 0 .25 0 0 .25   .25 0 .25 0 0 .25 .25   .25 0 0 .25 .25 0 .25   0
.25 .25 0 .25 0 .25   0 .25 0 0 .25 .25 .25   0 0 .25 .25 0 .25 .25   .142857 .142857
.142857 .142857 .142857 .142857 .142857;
let Y[7,1] = 50.7400 64.7800 52.2567 61.6100 45.4933 66.3433 56.4100;
b = inv(X′*X)*X′* (Y); b′;
var = inv(X′*X); var;
```

In the above, $.25 = \frac{1}{4}$ and $.142857 = \frac{1}{7}$, format 2,4 requests four decimals in the output, and b gives the cultivar means. The cultivar means for A to G, respectively, are given by b as 49.9316, 36.7494, 106.7805, 56.1849, 42.8161, 42.3138, and 60.0940. For this treatment design, it is necessary to use the actual coefficients of the mixture means rather than zeros and ones as was done in Example 16.1. In this form, the mixture means will all have the same variance, whereas if the first six equations were multiplied by 4 and the last equation by 7, the resulting entities in the Y vector would have unequal variances. The resulting output for b gives the above cultivar means, and var equals variance, which is

$$\begin{pmatrix} 15.5185 & 4.8519 & 4.8519 & -0.4815 & 4.8519 & 4.8519 & -18.1111 \\ 4.8519 & 15.5185 & 4.8519 & 4.8519 & -0.4815 & 4.8519 & -18.1111 \\ 4.8519 & 4.8519 & 15.5185 & 4.8519 & 4.8519 & -0.4815 & -18.1111 \\ -0.4815 & 4.8519 & 4.8519 & 15.5185 & 4.8519 & 4.8519 & -18.1111 \\ 4.8519 & -0.4815 & 4.8519 & 4.8519 & 15.5185 & 4.8519 & -18.1111 \\ 4.8519 & 4.8519 & -0.4815 & 4.8519 & 4.8519 & 15.5185 & -18.1111 \\ -18.1111 & -18.1111 & -18.1111 & -18.1111 & -18.1111 & -18.1111 & 59.6668 \end{pmatrix}.$$

The variance of an element in the Y vector is the error mean square divided by number of replicates = $27.2871/3 = 9.0957$. A variance of a difference between A and B means, say, is 9.0957 (15.5185 + 15.5185 - 4.8519 - 4.8519) = 194.0404; the standard error of the difference is $\sqrt{194.0404} = 13.93$. The variance of a difference between means A and D, say, is 9.0957(15.5185 + 15.5185 + 0.4815 + 0.4815) = 291.0624, with a standard error of a difference of 17.06. The variance of a difference of G and any other cultivar is 9.0957(15.5185 + 59.6668 + 18.1111 + 18.1111) = 1013.3292 and a standard error of a difference of 31.83.

The mean of the cultivar means, 56.4100, provides an estimate of $\mu + \delta.$, using the previous restriction on the solutions for the cultivar effects. The cultivar means minus 56.4100 results in solutions for the cultivar effects as follows:

$$\hat{\tau}_A + \hat{\delta}_A - \hat{\delta}. = 49.9316 - 56.4100 = -6.4784,$$

$$\hat{\tau}_B + \hat{\delta}_B - \hat{\delta}. = 36.7494 - 56.4100 = -19.6606,$$

$$\hat{\tau}_C + \hat{\delta}_C - \hat{\delta}_. = 106.7805 - 56.4100 = 50.3705,$$

$$\hat{\tau}_D + \hat{\delta}_D - \hat{\delta}_. = 56.1849 - 56.4100 = -0.2251,$$

$$\hat{\tau}_E + \hat{\delta}_E - \hat{\delta}_. = 42.8161 - 56.4100 = -13.5939,$$

$$\hat{\tau}_F + \hat{\delta}_F - \hat{\delta}_. = 42.3138 - 56.4100 = -14.0962,$$

$$\hat{\tau}_G + \hat{\delta}_G - \hat{\delta}_. = 60.0940 - 56.4100 = 3.6840.$$

These solutions add to zero within rounding error.

16.4 Response Model Equations, Solutions, and Analyses, Cultivar + BSMA Effects

Response model equations for estimating both cultivar and BSMA effects for $n = 2$, 3, and 4 are

$$Y_{ghi} = \mu + \rho_g + (\gamma_h + \gamma_i)/2 + \pi_{hi} + \epsilon_{ghi}, \tag{16.6}$$

$$Y_{ghij} = \mu + \rho_g + (\gamma_h + \gamma_i + \gamma_j)/3 \tag{16.7}$$

$$+ 2(\pi_{hi} + \pi_{hj} + \pi_{ij})/3 + \epsilon_{ghij}, \tag{16.8}$$

$$Y_{ghijk} = \mu + \rho_g + (\gamma_h + \gamma_i + \gamma_j + \lambda_k)/4$$

$$+ (\pi_{hi} + \pi_{hj} + \pi_{hk} + \pi_{ij} + \pi_{ik} + \pi_{jk})/2 + \epsilon_{ghijk},$$

where the π's are the BSMA effects and the other terms are as defined above. Since any pair, say hi, of cultivars for $n = 4$ only occupies one-half of the e.u. as compared to $n = 2$, the coefficient of one-half is used. Likewise for $n = 3$, a pair hi only occupies two-thirds of the e.u. and the coefficient 2/3 brings the BSMA effects to the same basis as for $n = 2$. In general, this coefficient is $2/n$. Extension to mixtures of size $n > 4$ is straightforward. The additional restrictions over those in Section 16.3 are that the sum of the BSMA effects for any cultivar, say h, is zero. The solution for the replicate effect is the same as in (16.1). If the TD is a BIBD, the solution for the hth cultivar effect is the last part of (16.5). If the TD is a BIBD for $n = 2$, with $v = m(m-1)/2$, the solution for the cultivar effect is

$$\hat{\gamma}_h = \frac{2(m-1)}{m-2} \left(\bar{y}_{\cdot h \cdot} - \bar{y}_{\dots}\right)$$

and the solution for a BSMA effect is

$$\hat{\pi}_{hi} = \bar{y}_{\cdot hi} + \frac{m}{m-2}\, \bar{y}_{\dots} - \frac{m-1}{m-2} \left(\bar{y}_{\cdot h \cdot} + \bar{y}_{\cdot\cdot i}\right).$$

The variances of a difference are

$$\text{Var}\left(\hat{\gamma}_h - \hat{\gamma}_{h'}\right) = \frac{4(m-1)^2}{(m-2)^2}\, \text{Var}\left(\bar{y}_{\cdot h \cdot} - \bar{y}_{\cdot h' \cdot}\right) = \frac{8\sigma_\epsilon^2}{r(m-2)}$$

and

$$\operatorname{Var}\left(\hat{\pi}_{hi} - \hat{\pi}_{hi'}\right) = \operatorname{Var}\left(\bar{y}_{\cdot hi} - \bar{y}_{\cdot hi'} - \frac{m-1}{m-2}\left(\bar{y}_{\cdot\cdot i} - \bar{y}_{\cdot\cdot i'}\right)\right) = \frac{2\sigma_\epsilon^2}{r}\frac{(m-3)}{(m-2)}.$$

For $n > 2$, the design is usually not balanced for both cultivar and BSMA effects. For this situation, the procedure described in the following example may be used to obtain variances of differences.

16.4.1 Example 16.3. Minimal Treatment Design for Cultivar and BSMA Effects

The treatment design for the data in Table 16.8 was obtained by deleting the combinations ABC, ABF, ADF, BCD, and CEF from the 20 mixtures obtained by taking all combinations of $m = 6$ cultivars taken $n = 3$ at a time. This set of 15 results in unequal occurrences of the 6 cultivars which may be considered undesirable but is a minimal saturated treatment design.

The $3(15) = 45$ responses given in Table 16.8 are used to illustrate the computations involved in obtaining an ANOVA, cultivar effects, and BSMA effects. An ANOVA for a standard RCBD is first obtained and given in the bottom part of Table 16.8. It would appear that there are some relatively large residuals which would alert the need to search for outliers in the data. This will not be done, as the calculational procedure is the goal here. A study of residuals is left as an exercise for the reader. From the results in the ANOVA, it appears that there are no significant differences among the 15 mixtures with the F-value for treatments being a little larger than 1.

Since this is a nonorthogonal design, use may be made of such software packages as MAPLE, GAUSS, MATHEMATICA, etc. to obtain an analysis. GAUSS is used here. A program used for the order of treatments in Table 16.8 is

```
let X[15,21] =
1 1 0 1 0 0   2 0 2 0 0   0 2 0 0   0 0 0   0 0 0
1 1 0 0 1 0   2 0 0 2 0   0 0 2 0   0 0 0   0 0 0
1 0 1 1 0 0   0 2 2 0 0   0 0 0 0   2 0 0   0 0 0
1 0 1 0 1 0   0 2 0 2 0   0 0 0 0   0 2 0   0 0 0
1 0 1 0 0 1   0 2 0 0 2   0 0 0 0   0 0 2   0 0 0
1 0 0 1 0 1   0 0 2 0 2   0 0 0 0   0 0 0   0 2 0
1 0 0 0 1 1   0 0 0 2 2   0 0 0 0   0 0 0   0 0 2
0 1 1 0 1 0   0 0 0 0 0   2 0 2 0   0 2 0   0 0 0
0 1 1 0 0 1   0 0 0 0 0   2 0 0 2   0 0 2   0 0 0
0 1 0 1 1 0   0 0 0 0 0   0 2 2 0   0 0 0   2 0 0
0 1 0 1 0 1   0 0 0 0 0   0 2 0 2   0 0 0   0 2 0
0 1 0 0 1 1   0 0 0 0 0   0 0 2 2   0 0 0   0 0 2
0 0 1 1 1 0   0 0 0 0 0   0 0 0 0   2 2 0   2 0 0
0 0 1 1 0 1   0 0 0 0 0   0 0 0 0   2 0 2   0 2 0
0 0 0 1 1 1   0 0 0 0 0   0 0 0 0   0 0 0   2 2 2;
```

TABLE 16.8. Barley Plant Weights and Minimal TD for Estimating Cultivar and BSMA Effects.

Mixture	Replicate 1	Replicate 2	Replicate 3	Mixture total	Mixture mean
ABD	36.04	46.61	37.60	120.25	40.0833
ABE	37.97	51.60	49.28	138.85	46.2833
ACD	34.12	42.43	56.54	133.09	44.3633
ACE	43.17	32.23	40.33	115.73	38.5767
ACF	38.32	39.69	38.27	116.28	38.7600
ADF	42.50	47.24	48.98	138.72	46.2400
AEF	33.31	39.68	43.37	116.36	38.7867
BCE	31.78	34.39	37.72	103.89	34.6300
BCF	34.19	45.61	55.18	134.98	44.9933
BDE	41.88	41.21	49.18	132.27	44.0900
BDF	40.69	60.72	46.91	148.32	40.4400
BEF	43.17	36.70	44.69	124.56	41.5200
CDE	34.39	54.95	34.61	123.95	41.3167
CDF	43.84	34.55	48.75	127.14	42.3800
DEF	39.36	48.37	42.93	130.66	43.5533
Total	574.33	655.98	674.34	1,905.05	42.3344

ANOVA

Source of variation	d.f.	SS	MS
Total	45	82,773.6861	
Correction for mean	1	80,649.2334	
Replicate $= R$	2	374.6845	187.3423
Treatment $= T$	14	608.1038	43.4360
$R \times T$	28	1,141.6644	40.7737
Cultivar	5	138.7791	27.7558
BSMA	9	469.3263	52.1474

```
let Y[15,1] = 120.25 138.85 133.09 115.73 116.28 138.72 116.36 103.89 134.98
132.27 148.32 124.56 123.95 127.14 130.66;
S = ones(1,15); S*Y; let J2[6,15] =
1 1 1 1 1   0 0 0 0   0 0 0   0 0 0
1 0 0 0 0   1 1 1 1   0 0 0   0 0 0
0 1 0 0 0   1 0 0 0   1 1 1   0 0 0
0 0 1 0 0   0 1 0 0   1 0 0   1 1 0
0 0 0 1 0   0 0 1 0   0 1 0   1 0 1
0 0 0 0 1   0 0 0 1   0 0 1   0 1 1;
J1 = zeros(21,6); J3 = zeros(15,15); J = J1~(J2:J3);n = 3; r = 3;
var = inv(X' * X - J); var; b = var*X' * Y/n; b'; X' * Y/n; Y' * Y/r; b' * (X' * Y/n);
```

The zeros and ones in the X matrix correspond to the numerators of the fractions in equation (16.7) for the particular effect. The effects are listed in alphabetical order. The entries in Y are n/r times the mixture totals. $S^{*}Y$ for $n = r$ gives the totals for the experiment and is a check on the entries in Y. The J matrix is used to add the restrictions that the sum of the interaction effects for each cultivar is zero. b gives the solutions for the cultivar means and BSMA effects. $b'^{*}X'^{*}Y/n = Y'^{*}Y/r$ for this example since n is constant for every mixture. n^2/r times the error mean square is the constant multiplier for var to obtain the variance–covariance matrix for the cultivar means and BSMA effects. The resulting 21-by-21 variance–covariance matrix is not reproduced here owing to its size but may be obtained from the above GAUSS program.

The cultivar means are

$$\hat{\mu} + \hat{\tau}_A + \hat{\delta}_A = 50.6860,$$
$$\hat{\mu} + \hat{\tau}_B + \hat{\delta}_B = 49.5660,$$
$$\hat{\mu} + \hat{\tau}_C + \hat{\delta}_C = 27.4160,$$
$$\hat{\mu} + \hat{\tau}_D + \hat{\delta}_D = 52.8260,$$
$$\hat{\mu} + \hat{\tau}_E + \hat{\delta}_E = 32.2660,$$
$$\hat{\mu} + \hat{\tau}_F + \hat{\delta}_F = 40.5660.$$

The BSMA effects from the above GAUSS program are

$$\hat{\pi}_{AB} = -2.5130, \qquad \hat{\pi}_{BF} = +9.0995,$$
$$\hat{\pi}_{AC} = +2.5345, \qquad \hat{\pi}_{CD} = +3.1945,$$
$$\hat{\pi}_{AD} = -4.6480, \qquad \hat{\pi}_{CE} = -5.3455,$$
$$\hat{\pi}_{AE} = +5.4920, \qquad \hat{\pi}_{CF} = -2.8630,$$
$$\hat{\pi}_{AF} = -0.8655, \qquad \hat{\pi}_{DE} = +7.8720,$$
$$\hat{\pi}_{BC} = +2.4795, \qquad \hat{\pi}_{DF} = +2.8345,$$
$$\hat{\pi}_{BD} = -9.2530, \qquad \hat{\pi}_{EF} = -8.2055.$$
$$\hat{\pi}_{BE} = +0.1870,$$

Note that the sum of the BSMA effects for a cultivar equals zero. The mean of the cultivar means is 42.2210, which differs from the mean of the 45 reponses, 42.3344, since there are unequal ocurrences of the cultivars in the mixures, but owing to the small cultivar effects, these means differ little. The cultivar effects are obtained by subtracting 42.2210 from each of the cultivar means. An estimate of $\mu + \delta.$ is 42.2210. A variance of a difference between two cultivar means is $\left[(3^2(40.7737/3)\right](0.9000 + 0.9000 - (-0.1000 - 0.1000)) = 244.6422$, and a standard error of a difference of 15.64. Variances of differences and standard errors of a difference between BSMA effects vary with the effects and may be obtained from the program output for var.

16.5 Response Model Equations, Solutions, and Analyses, Cultivar + BSMA + TSMA Effects

Response models for estimating cultivar, BSMA, and TSMA effects for $n = 3$ and 4 are

$$\begin{aligned} Y_{ghij} &= \mu + \rho_g + (\gamma_h + \gamma_i + \gamma_j)/3 + 2(\pi_{hi} + \pi_{hj} + \pi_{ij})/3 \\ &\quad + \eta_{hij} + \epsilon_{ghij}, \end{aligned} \tag{16.9}$$

$$\begin{aligned} Y_{ghijk} &= \mu + \rho_g + (\gamma_h + \gamma_i + \gamma_j + \lambda_k)/4 \\ &\quad + (\pi_{hi} + \pi_{hj} + \pi_{hk} + \pi_{ij} + \pi_{ik} + \pi_{jk})/2 \\ &\quad + 3(\eta_{hij} + \eta_{hik} + \eta_{hik} + \eta_{ijk})/4 + \epsilon_{ghijk}, \end{aligned} \tag{16.10}$$

where η_{hij} is the TSMA effect for cultivars h, i, and j and the other symbols are as defined above. The coefficient 3/4 is used to place the TSMA effect on the same basis as other effects since only 3/4 of the e.u. is available for any one of these effects.

16.5.1 *Example 16.4. Estimating Cultivar, BSMA, and TSMA Effects*

The data for illustrating the computations for obtaining solutions for cultivar means and effects, BSMA effects, and TSMA effects are given in Table 16.9. The responses are for biomass of cultivars in the mixture. The TD is all possible combinations of the $m = 6$ cultivars in mixtures of $n = 3$, i.e., $v = 20$ mixtures. A standard RCBD ANOVA for the $r = 3$ replicates and $v = 20$ mixtures is given in the bottom part of Table 16.9. Using response equation (16.10) and a GAUSS program (see Appendix 16.2), similar to the one in Example 16.3 with appropriate changes in the X and Y matrices, the various means and solutions for effects are obtained. The means $\bar{y}_{\cdot h\cdot\cdot}$ are obtained as follows:

$$\begin{aligned} \bar{y}_{\cdot A\cdot\cdot} &= (68.50 + 31.97 + 33.05 + 20.90 + 62.76 + 72.54 + 78.06 \\ &\quad + 30.03 + 24.46 + 42.26)/30 = 15.4843, \\ \bar{y}_{\cdot B\cdot\cdot} &= 12.0713, \\ \bar{y}_{\cdot C\cdot\cdot} &= 21.8907, \\ \bar{y}_{\cdot D\cdot\cdot} &= 11.6963, \\ \bar{y}_{\cdot E\cdot\cdot} &= 14.1233, \\ \bar{y}_{\cdot F\cdot\cdot} &= 12.8650. \end{aligned}$$

The means $\bar{y}_{\cdot h\cdot\cdot}$, $\bar{y}_{\cdot hi\cdot}$, and $\hat{y}_{\cdot hij}$ may be obtained by setting up a diagonal matrix of the reciprocals of the number of mixtures in each of the means times the coefficient in the X matrix; i.e., 0.1, 0.125, and 0.33333 are the elements of the diagonal matrix.

TABLE 16.9. Biomass for $v = 20$ Mixtures of $n = 3$ Cultivars of $m = 6$ Cultivars to Obtain Cultivar Means and Effects, BSMA Effects, and TSMA Effects with an ANOVA.

Mixture	Replicate 1	Replicate 2	Replicate 3	Mixture total	Mixture mean
ABC	36.02	22.34	10.14	68.50	22.8333
ABD	8.37	10.58	13.02	31.97	10.6567
ABE	9.83	13.07	10.15	33.05	11.0167
ABF	6.20	6.31	8.39	20.90	6.9667
ACD	22.27	17.56	22.93	62.76	20.9200
ACE	18.90	32.93	20.71	72.54	24.1800
ACF	33.47	17.85	26.74	78.06	26.0200
ADE	11.28	8.32	10.43	30.03	10.0100
ADF	11.64	3.25	9.57	24.46	8.1533
AEF	12.77	12.41	17.08	42.26	14.0867
BCD	14.51	11.69	11.26	37.46	12.4867
BCE	26.78	28.92	25.24	80.94	26.9800
BDE	23.05	16.24	19.69	59.08	19.6933
BDF	4.19	9.49	2.11	15.79	5.2633
BDF	1.16	0.55	0.82	2.53	0.8433
BEF	2.37	6.05	3.50	11.92	3.9733
CDE	22.69	15.92	23.71	62.32	20.7733
CDF	15.89	35.98	20.02	71.89	23.9633
CEF	23.40	15.17	24.60	63.17	21.0567
DEF	6.38	2.17	3.13	11.68	3.8933
Total	311.17	286.90	283.24	881.31	14.6885

ANOVA

Source of variation	d.f.	SS	MS
Total	60	18,011.9325	
Correction for mean	1	12,945.1219	
Replicate $= R$	2	23.0419	
Mixture $= M$	19	4,003.3098	210.7005
$R \times M$	38	1,040.4589	27.3805
Cultivar	5	3,597.5998	719.5200
BSMA	9	195.5267	21.7252
TSMA	5	210.1958	42.0392

Denote this matrix as DR. Then $(X'Y/r)^*$ DR produces these means where $\bar{y}_{\cdot hij}$ are the means in the last column of Table 16.9. The $\bar{y}_{\cdot hi\cdot}$ means are obtained as

$$\bar{y}_{\cdot AB\cdot} = 12.8683, \quad \bar{y}_{\cdot BC\cdot} = 20.4983, \quad \bar{y}_{\cdot CE\cdot} = 23.2475,$$

$$\bar{y}_{\cdot AC\cdot} = 23.4833, \quad \bar{y}_{\cdot BD\cdot} = 7.3125, \quad \bar{y}_{\cdot CF\cdot} = 22.6833,$$

$$\bar{y}_{\cdot AD\cdot} = 12.4350, \quad \bar{y}_{\cdot BE\cdot} = 11.8083, \quad \bar{y}_{\cdot DE\cdot} = 9.9850,$$

$$\bar{y}_{\cdot AE\cdot} = 14.8233, \quad \bar{y}_{\cdot BF\cdot} = 7.8692, \quad \bar{y}_{\cdot DF\cdot} = 9.2133,$$

$$\bar{y}_{\cdot AF\cdot} = 13.8067, \quad \bar{y}_{\cdot CD\cdot} = 19.5358, \quad \bar{y}_{\cdot EF\cdot} = 10.7525.$$

The cultivar means obtained from the above program are

$$\begin{aligned}
\hat{\mu} + \hat{\tau}_A + \hat{\delta}_A &= 18.6667,\\
\hat{\mu} + \hat{\tau}_B + \hat{\delta}_B &= 1.6027,\\
\hat{\mu} + \hat{\tau}_C + \hat{\delta}_C &= 50.6993,\\
\hat{\mu} + \hat{\tau}_D + \hat{\delta}_D &= -0.2723,\\
\hat{\mu} + \hat{\tau}_E + \hat{\delta}_E &= 11.8627,\\
\hat{\mu} + \hat{\tau}_F + \hat{\delta}_F &= 5.5710.
\end{aligned}$$

The solution for $\mu + \delta.$ is the mean of the above 6 means, or 14.6885, or the mean of the 60 responses in Table 16.9. These means are equal since this is a BIBD. Following the procedure given by Federer and Raghavarao (1987), the cultivar effects, for $3(m-1)/(m-3) = 5$, are

$$\begin{aligned}
\hat{\tau}_A + \hat{\delta}_A - \hat{\delta}. &= 5(\bar{y}_{\cdot A\cdot\cdot} - \bar{y}_{\cdots}) = 5(15.4843 - 14.6885) = 3.9790,\\
\hat{\tau}_B + \hat{\delta}_B - \hat{\delta}. &= 5(\bar{y}_{\cdot B\cdot\cdot} - \bar{y}_{\cdots}) = 5(12.0713 - 14.6885) = -13.0860,\\
\hat{\tau}_C + \hat{\delta}_C - \hat{\delta}. &= 5(\bar{y}_{\cdot C\cdot\cdot} - \bar{y}_{\cdots}) = 5(21.8907 - 14.6885) = 36.0110,\\
\hat{\tau}_D + \hat{\delta}_D - \hat{\delta}. &= 5(\bar{y}_{\cdot D\cdot\cdot} - \bar{y}_{\cdots}) = 5(11.6963 - 14.6885) = -14.9610,\\
\hat{\tau}_E + \hat{\delta}_E - \hat{\delta}. &= 5(\bar{y}_{\cdot E\cdot\cdot} - \bar{y}_{\cdots}) = 5(14.1233 - 14.6885) = -\ 2.8260,\\
\hat{\tau}_F + \hat{\delta}_F - \hat{\delta}. &= 5(\bar{y}_{\cdot F\cdot\cdot} - \bar{y}_{\cdots}) = 5(12.8650 - 14.6885) = -\ 9.1175.
\end{aligned}$$

The solutions obtained by subtracting the mean of the cultivar means for A, B, C, D, E, and F, respectively, are $18.6677 - 14.6885 = 3.9792$, -13.0858, 36.0108, -14.9608, -2.8258, and -9.1175. These values agree with the above solutions within rounding errors. The solutions obtained from the GAUSS program for BSMA effects are

$$\begin{aligned}
\hat{\pi}_{AB} &= +1.3695, & \hat{\pi}_{BC} &= +0.2358, & \hat{\pi}_{CE} &= +0.7882,\\
\hat{\pi}_{AC} &= -3.5930, & \hat{\pi}_{BD} &= -1.0930, & \hat{\pi}_{CF} &= +3.8145,\\
\hat{\pi}_{AD} &= +1.4758, & \hat{\pi}_{BE} &= +3.2932, & \hat{\pi}_{DE} &= -0.7705,\\
\hat{\pi}_{AE} &= -0.4605, & \hat{\pi}_{BF} &= -3.8055, & \hat{\pi}_{DF} &= +1.6332,\\
\hat{\pi}_{AF} &= +1.2083, & \hat{\pi}_{CD} &= -1.2455, & \hat{\pi}_{EF} &= -2.8505.
\end{aligned}$$

The solutions obtained for TSMA effects are (lowercase letters for subscripts are used for clarity)

$$\hat{\eta}_{abc} = +0.5019, \quad \hat{\eta}_{ade} = -0.2392, \quad \hat{\eta}_{bdf} = +0.7197,$$

$$\hat{\eta}_{abd} = +2.8225, \quad \hat{\eta}_{adf} = -2.7136, \quad \hat{\eta}_{bef} = -0.1303,$$

$$\hat{\eta}_{abe} = -2.4958, \quad \hat{\eta}_{aef} = +3.4547, \quad \hat{\eta}_{cde} = +0.8286,$$

$$\hat{\eta}_{abf} = -0.8286, \quad \hat{\eta}_{bcd} = -3.4547, \quad \hat{\eta}_{cdf} = +2.4958,$$

$$\hat{\eta}_{acd} = +0.1303, \quad \hat{\eta}_{bce} = +2.7136, \quad \hat{\eta}_{cef} = -2.8225,$$

$$\hat{\eta}_{ace} = -0.7197, \quad \hat{\eta}_{bcf} = +0.2392, \quad \hat{\eta}_{def} = -0.5019,$$

$$\hat{\eta}_{acf} = +0.0875, \quad \hat{\eta}_{bde} = -0.0875,$$

The sum of squares for mixtures is obtained from the GAUSS program as $Y'Y/r$ minus the correction for mean. Let b equal the solution for the set of effects and X be the design matrix. Then $b'X'Y/r$ is equal to $Y'Y/r$ as this is standard regression theory in that the regression coefficient times the sum of the cross products is the sum of squares due to regression. The mixture sum of squares may be partitioned as follows for this design. The sum of squares due to cultivar effects is the sum of the products of cultivar effects and the sum of the mixture means in which the cultivar occurred. It is equal to

$$3.9792(154.8433) - 13.0858(120.7133) + 36.0108(218.9067)$$
$$- 14.9608(116.9633) - 2.8258(141.2333) - 9.1175(128.6500) = 3,597.5998.$$

The sum of squares for BSMA effects is obtained as the sum of the products of 2 times the sum of the mixture means in which the pair of cultivars hi occurred times the BSMA effect. This sum of squares is

$$102.9467(1.3695) + 187.9067(-3.5930) + 99.4800(1.4758)$$
$$+ 118.5867(-0.4605) + 110.4533(1.2083) + 163.9867(0.2358)$$
$$+ 58.5000(-1.0930) + 94.4667(3.2932) + 62.9533(-3.8055)$$
$$+ 156.2867(-1.2455) + 185.9800(0.7882) + 181.4667(3.8145)$$
$$+ 79.8800(-0.7705) + 73.7067(1.6332) + 86.0200(-2.8505) = 195.5267.$$

The sum of squares for TSMA effects is computed as the sum of the products of n times the mixture mean times the TSMA effect and is

$$68.5000(0.5019) + 31.9700(2.8225) + 33.0500(-2.4958)$$
$$+ 20.9000(-0.8286) + 62.7600(0.1303) + 72.5400(-0.7197)$$
$$+ 78.0600(0.0875) + 30.0300(-0.2392) + 24.4600(-2.7136)$$
$$+ 42.2600(3.4547) + 37.4600(-3.4547) + 80.9400(2.7136)$$

$$+ 59.0800(0.2392) + 15.7900(-0.0875) + 2.5300(0.7197)$$

$$+ 11.9200(-0.1303) + 62.3200(0.8286) + 71.8900(2.4958)$$

$$+ 63.1700(-2.8225) + 11.6800(-0.5019) = 210.1958.$$

From the partitioning of the mixture sum of squares, it is seen that the major part is attributable to the cultivar effects, BSMA effects appear to be nonexistent, and there is small indication of TSMA effects as $F = 1.54$ (approximately at the 0.2 level). The variance–covariance matrix may be obtained from the GAUSS output denoted as var. Since the variance of the elements of the Y vector is $n^2\hat{\sigma}_\epsilon^2/r$, 9(27.3805)/3 times var results in the variance–covariance matrix for the 6 + 15 + 20 = 41 effects.

Variances of differences between effects may be obtained from the GAUSS output. The variance of a difference between two cultivar effects is

$$3(27.3805)(0.9000 + 0.9000 - (-0.1000 - 0.1000)) = 164.2830,$$

and a standard error of a difference of 12.82. The variance of a difference between two BSMA effects, e.g., AB and AC, is

$$3(27.3805)(0.4500 + 0.4500 - (-0.1125 - 0.1125)) = 92.4092,$$

and a standard error of a difference of 9.61. The variance of a difference between two TSMA effects, e.g., ABC and ABD, is

$$3(27.3805)(0.5278 + 0.5278 - (0.0463 + 0.0463)) = 79.1023,$$

and standard error of 8.89.

16.6 Response Model Equations, Solutions, and Analyses, Cultivar+BSMA+ TSMA+QSMA Effects

Response model equations for estimating cultivar, BSMA, TSMA, and QSMA effects for $n = 4$ and 5 are

$$\begin{aligned}
Y_{ghijk} &= \mu + \rho_g + (\gamma_h + \gamma_i + \gamma_j + \gamma_k)/4 \\
&\quad + (\pi_{hi} + \pi_{hj} + \pi_{hk} + \pi_{ij} + \pi_{ik} + \pi_{jk})/2 \\
&\quad + 3(\eta_{hij} + \eta_{hik} + \eta_{hjk} + \eta_{ijk})/4 + \psi_{hijk} + \epsilon_{ghijk}, \qquad (16.11) \\
Y_{ghijkl} &= \mu + \rho_g + (\gamma_h + \gamma_i + \gamma_j + \gamma_k + \gamma_l)/5 \\
&\quad + 2(\pi_{hi} + \pi_{hj} + \pi_{hk} + \pi_{hl} + \pi_{ij} + \pi_{ik} + \pi_{il} + \pi_{jk} \\
&\quad + \pi_{jl} + \pi_{kl})/5 + 3(\eta_{hij} + \eta_{hik} + \eta_{hil} + \eta_{hjk} + \eta_{hjl} \\
&\quad + \eta_{hkl} + \eta_{ijk} + \eta_{ijl} + \eta_{ikl} + \eta_{jkl})/5 + 4(\psi_{hijk} + \psi_{hijl}
\end{aligned}$$

$$+ \psi_{hikl} + \psi_{hjkl} + \psi_{ijkl})/5 + \phi_{hijkl} + \epsilon_{ghijkl}, \tag{16.12}$$

where ψ_{hijk} is the QSMA effect for cultivars h, i, j, and k, ϕ_{hijkl} is a specific mixing effect among the five cultivars h, i, j, k, and l, and the other symbols are as defined previously. The coefficients 1/4, 1/2, and 3/4 and the coefficients 1/5, 2/5, 3/5, and 4/5 put the effects on the same basis as described previously for $n = 4$ and $n = 5$ cultivars in a mixture, respectively. Extensions of the above response model equations to $n > 5$ cultivars in a mixture is straightforward. When the number of combinations is $v = m!/n!(m-n)!$, where ! denotes factorial, the results of Federer and Raghavarao (1987) may be used to obtain solutions for the various effects. The computer programs described previously may also be used for the numerical analysis of data from these mixture experiments.

16.7 Response Models for Crop Competition

Competition studies among plant species and cultivars are commonplace in the literature. Absence of competition between experimental units is important in agronomic and plant breeding experiments in order to obtain unbiased estimates of responses. Various statistical procedures have been advanced to summarize results from competition experiments; for example, two co-authored papers on such procedures are Jensen and Federer (1965) and Federer and Basford (1991). In addition, four competition models for pairs of crops were proposed in Volume I, Section 7.4. These models are generalized in the present section to include additional cultivars. Treatment designs balanced for competition effects are given in Section 8.2 of Volume I. Other TDs are possible. For example, consider the following set of experimental units for $m = 4$ cultivars, where the center two letters represent cultivar $h = a$, and a, b, c, and d are the bordering cultivars:

$$\begin{array}{ccccccc} aaaa & baab & caac & daad & baac & baad & caad \\ aaaa & baab & caac & daad & baac & baad & caad \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ aaaa & baab & caac & daad & baac & baad & caad \end{array}$$

Cultivar a is bordered by itself, by b, by c, by d, by b and c, by b and d, and by c and d. The center two columns are used to obtain the responses for cultivar a. A treatment design would be to use an additional seven e.u.s for each of the cultivars b, c, and d, or the above seven might suffice if interest centers solely on cultivar a. Instead of placing a cultivar on each side of cultivar a, a mixture of two (or more) cultivars could be used on each side of the specified cultivar. For border cultivars b, c, d, and e, each side of cultivar a could be bordered by the mixture $bcde$, by mixture bc on one side and mixture de on the other, or by b as one border and mixture cde as the other. As is obvious, many TD arrangements are possible, which means that the experimenter needs to select the correct TD meeting the goals of the experiment.

16.7.1 Model 1 for Crop Competition

The TD described above is a generalization of the design used by Jensen and Federer (1965). Response models for sole crop h (cultivar bordered by itself) and for h bordered by another cultivar is presented in Section 7.4 of Volume I. Given that the experiment design is a randomized complete block (RCBD), response models for a sole crop and crop h with a border of two cultivars are

$$Y_{gh} = \mu + \rho_g + \tau_h + \epsilon_{gh},$$

$$Y_{gh(ij)} = \mu + \rho_g + \tau_h + \delta_h + (\gamma_{h(i)} + \gamma_{h(j)})/2 + \pi_{h(ij)} + \epsilon_{gh(ij)},$$

where the first equation represents the mean effect, the block effect, cultivar h sole crop effect, and a random error; in the second equation, δ_h is the effect on cultivar h by all the bordering mixtures, $\gamma_{h(i)}$ is the border effect on cultivar h response by cultivar i, $\pi_{h(ij)}$ is the interaction border effect of the combination or mixture of cultivars i and j on cultivar h response, and the epsilon terms represent random error terms with variance σ_ϵ^2. Consider the case where a cultivar h is bordered by two different cultivars i and j. Using the restriction that the sum of the block effects, the sum of the border effects summed over $h \neq i = 1, 2, \ldots, m$ is zero, and the sum of the interaction border effects summed over $h, i \neq j = 1, 2, \ldots, m$ is zero, solutions for the various effects are

$$\mu = \bar{y}_{\cdot\cdot} = \sum_{g=1}^{r} \sum_{h=1}^{m} \frac{Y_{gh}}{rm},$$

$$\hat{\tau}_h = \bar{y}_{\cdot h} - \bar{y}_{\cdot\cdot},$$

$$\hat{\delta}_h = \bar{y}_{\cdot h(\cdot\cdot)} - \bar{y}_{\cdot h},$$

$$\hat{\gamma}_{h(i)}/2 = \frac{m-2}{m-3}\left(\bar{y}_{\cdot h(i\cdot)} - \bar{y}_{\cdot h(\cdot\cdot)}\right),$$

$$\hat{\gamma}_{h(j)}/2 = \frac{m-2}{m-3}\left(\bar{y}_{\cdot h(\cdot j)} - \bar{y}_{gh(\cdot\cdot)}\right),$$

$$\hat{\pi}_{h(ij)} = \bar{y}_{\cdot h(ij)} - \frac{m-2}{m-3}\left(\bar{y}_{\cdot h(i\cdot)} + \bar{y}_{\cdot h(\cdot j)}\right) + \frac{m-1}{m-3}\,\bar{y}_{\cdot h(\cdot\cdot)}.$$

Variances for differences of the various effects are

$$\mathrm{Var}(\hat{\tau}_h - \hat{\tau}_{h'}) = \mathrm{Var}(\bar{y}_{\cdot h} - \bar{y}_{\cdot h'}) = \frac{2\sigma_\epsilon^2}{r}, \quad h \neq h',$$

$$\mathrm{Var}(\hat{\delta}_h - \hat{\delta}_{h'}) = \mathrm{Var}\left(\bar{y}_{\cdot h(\cdot\cdot)} - \bar{y}_{\cdot h'(\cdot\cdot)}\right)$$

$$= \frac{4\sigma_\epsilon^2}{r(m-1)(m-2)}, \quad h \neq h',$$

$$
\begin{aligned}
\operatorname{Var}\left(\hat{\gamma}_{h(i)} - \hat{\gamma}_{h(i')}\right) &= 4\left(\frac{m-2}{m-3}\right)^2 \operatorname{Var}\left(\bar{y}_{\cdot h(i\cdot)} - \bar{y}_{\cdot h(i'\cdot)}\right) \\
&= \frac{2\sigma_\epsilon^2}{r(m-3)}, \quad i \neq i', \\
\operatorname{Var}\left(\hat{\pi}_{h(ij)} - \hat{\pi}_{h(ij')}\right) &= \operatorname{Var}\left(\bar{y}_{\cdot h(ij)} - \bar{y}_{\cdot h(ij')},\right. \\
&\quad \left. - \frac{m-2}{m-3}\left(\bar{y}_{\cdot h(\cdot j)} - \bar{y}_{\cdot h(\cdot j\prime)}\right)\right) = \frac{2(m-4)\sigma_\epsilon^2}{m-3}, \\
&\quad j \neq j'.
\end{aligned}
$$

The extension to more than two cultivars used as borders for cultivar h is straightforward. Since the algebraic expressions increase in complexity with the number of cultivars used simultaneously as borders, packages like MAPLE or MATHEMATICA will be useful in obtaining algebraic expressions for effects and their variances. The program given in Appendix 16.3 was used to obtain the above results.

16.7.2 Model 2 for Crop Competition

Model 2 for competition in Volume I was for the competition treatment design

aaa	*bab*	*cac*	*dad*
aba	*bbb*	*cbc*	*dbd*
aca	*bcb*	*ccc*	*dcd*
ada	*bdb*	*cdc*	*ddd*

For this TD, a standard two-factor factorial response model was presented. Even if the treament combinations

aab	*aac*	*aad*
abb	*cbb*	*dbb*
acc	*bcc*	*dcc*
add	*bdd*	*cdd*

were added to the previous set, the TD would not be a factorial arrangement for three factors. Instead, these 28 combinations would be all possible combinations of zero (cultivar bordered by itself), one, and two cultivars as borders. For square experimental units, there could be zero, one, two, three, or four cultivars used as borders. Hence, it appears that Model 2 usually is not extendable to TDs, which could be considered as three-factor, four-factor, etc. factorial models for crops and borders of crops. However, if number of borders was a factor, then number of borders, crop as a border, and crop bordered by another crop with this crop responses as the variable of interest, then this would be a three-factor type of experiment and could be treated as a factorial model. Another type of study would be to repeat the above designs over time, say, years, using a latin cube arrangement (see, e.g., Federer, 1967) to examine the effect of various borders in a rotation type of experiment. One possibility for a competition study with more than two factors is to consider three species with lines from each species in all combinations of crops and borders of crops.

An example of a three-factor intercropping experiment could be for cultivars of differing heights with rows oriented in an east-west direction. It could be that a shading border causes (a border on the north side) a different border effect than a border on the unshaded side (the south side). Or, it could be that rows oriented in a north-south direction would exhibit a morning sun effect (east side) and an afternoon sun effect (west side) as the border effect is exhibited. For such cultivars and effects, a three-factor factorial response model from a RCBD is

$$Yghij = \mu + \rho_g + \alpha_h + \beta_i + \gamma_j + \alpha\beta_{hi} + \alpha\gamma_{hj}$$

$$+ \beta\gamma_{ij} + \alpha\beta\gamma_{hij} + \epsilon_{ghij},$$

where μ is a common mean effect, ρ_g is the gth complete block effect, α_h is the effect of the hth cultivar being bordered, β_i is the border effect of cultivar i as a north border on crop h, γ_j is the border effect of cultivar j as a south border on cultivar h response, $\alpha\beta_{hi}$ is an interaction effect of cultivar h and border i, $\alpha\gamma_{hj}$ is an interaction effect of cultivar h and border j, $\beta\gamma_{hi}$ is an interaction effect of cultivars i and j as borders of cultivar k, $\alpha\beta\gamma_{hij}$ is a three-factor interaction of cultivar k bordered by cultivars i and j, and ϵ_{ghij} is a random error term distributed with mean zero and variance σ_ϵ^2. The treatment design for the above response model for m cultivars would be a 3^m factorial.

Another situation wherein the above factorial model would be appropriate is in the study of cultivars grown in sequences, i.e., first, second, and third in a sequence. There would be 3^m possible combinations for m cultivars.

16.7.3 *Model 3 for Crop Competition*

Model 3 in Volume I made use of the ideas of Martin (1980) to partition the factorial interaction terms into three component parts. Since factorial set-ups for competition studies for more than crops and borders do not appear to be practical, Model 3 is not generalized. A generalization for three factors, say, could be to partition each of the three two-factor interactions in the manner described for two factors and then determine if the three factor interaction can similarly be partitioned.

If a 3^m factorial design of the nature described for Model 2 is used, then extending the ideas of Martin (1980), we note that *each* two-factor interaction may be written as described for Model 3 in Volume I. This (from Volume I) is

$$\alpha\beta_{hi} = \eta_{hi\cdot} + \omega_{hi\cdot} + \kappa_{hi\cdot},$$

$$\alpha\gamma_{hj} = \eta_{h\cdot j} + \omega_{h\cdot j} + \kappa_{h\cdot j},$$

$$\beta\gamma_{ij} = \eta_{\cdot ij} + \omega_{\cdot ij} + \kappa_{\cdot ij},$$

where $\eta_{hi\cdot}$ is $\eta_{\alpha\beta}$ for $h = i$ and $-\eta_{\alpha\beta}/(m-1)$ for $h \neq i$,

$$\omega_{hi\cdot} = \omega_{ih\cdot} \quad \text{and} \quad \sum_{h \text{ or } i} \omega_{hi\cdot} = 0,$$

$$\kappa_{hi\cdot} = -\kappa_{ih\cdot} \quad \text{and} \quad \sum_{h \text{ or } i} \kappa_{hi\cdot} = 0,$$

and $\sum_h \alpha\beta_{hi} = \sum_i \alpha\beta_{hi} = 0$. The other two-factor interactions are defined similarly.

The three-factor interaction term may be partitioned for *each h* as

$$\alpha\beta\gamma_{hij} = \eta_{hij} + \omega_{hij} + \kappa_{hij},$$

where the η_{hij}, ω_{hij}, and κ_{hij} have the definitions given for two-factor interactions, but for each level of h, the cultivar whose response is measured in the center row.

16.7.4 Model 4 for Crop Competition

For the three-factor factorial described for Models 2 and 3, Model 4 of Volume I is easily extended to three factors by setting $\alpha\beta_{hh}$, $\alpha\gamma_{ii}$, and $\beta\gamma_{jj} = 0$ and having the sum of the remaining interaction effects sum to zero. Then, within these limitations, ω' and κ' effects are defined for each of the two-factor interactions and the three-factor interaction.

Model 4 for crop competition studies is similar to Model 3 and omits the combinations of crops bordered by themselves to compute the interaction terms, as this can radically change the size of the interaction terms. For any generalizations obtained for Model 3, these are easily extendable to Model 4.

16.8 Problems

16.1 A numerical example is constructed with known parameters with the TD described for Example 16.1. The purpose of this example is to demonstrate that the solutions for the parameters are the same as those used to construct the example. This device is useful in checking algebraic results. A check on sums of squares may also be made with such an example. For the example with $m = 6$ cultivars in mixtures of size $n = 3$, let the parameter values be

$$\begin{array}{lll}
\mu = 10, & \gamma_b = -3, & \epsilon_{2acf} = -1, \\
\rho_1 = -5, & \gamma_c = 0, & \epsilon_{1abd} = -1, \\
\rho_2 = 5, & \gamma_d = 0, & \epsilon_{1acf} = 1, \\
\gamma_a = -3, & \gamma_e = 3, & \epsilon_{2abd} = 1, \\
 & \gamma_f = 3, & \text{all other } \epsilon_{ghij} = 0.
\end{array}$$

For this set of parameters, note that the cultivar effects, $\gamma_h = \tau_h + \delta_h - \delta_.$, sum to zero and that the replicate by mixture sum of squares is 4 as there are 4 ± 1s to square and sum. The replicate sum of squares is $(m = 6)((-5)^2 + 5^2) = 300$, and the mixture sum of squares is $(r = 2)((-3)^2 + (-3)^2 + 3^2 + 3^2) = 72$. Using the above parameter values, the responses Y_{ghij} for replicate one are constructed as

$$Y_{1abd} = 10 - 5 + (-3 - 3 + 0)/3 - 1 = 2,$$

$$Y_{1acf} = 10 - 5 + (-3 + 0 + 3)/3 + 1 = 6,$$

$$Y_{1ade} = 10 - 5 + (-3 + 0 + 3)/3 + 0 = 5,$$

$$Y_{1bce} = 10 - 5 + (-3 + 0 + 3)/3 + 0 = 5,$$

$$Y_{1bef} = 10 - 5 + (-3 + 3 + 3)/3 + 0 = 6,$$

$$Y_{1cdf} = 10 - 5 + (0 + 0 + 3)/3 + 0 = 6.$$

Construction of the values for replicate 2 is left as an exercise for the reader. Carry out the calculations for cultivar effects as described in Example 16.1 and ascertain that the solutions are those used to construct the example.

16.2 Given the following data sets for plant biomass:

Mixture	Replicate 1	Replicate 2	Replicate 3
ABE	36.02	22.34	10.14
ACD	6.20	6.31	8.39
ADF	12.27	12.41	17.08
BCD	14.51	17.56	11.26
CDE	22.69	15.92	23.71
DEF	6.38	2.17	3.13

and

Mixture	Replicate 1	Replicate 2	Replicate 3
ABE	9.83	13.07	10.15
ACD	22.27	17.56	22.93
ADF	11.64	3.25	9.57
BCF	23.05	16.34	19.69
BDE	4.19	9.49	2.11
CEF	23.40	15.17	24.60

Obtain solutions for values for each of the two data sets and compare their values with those obtained from Example 16.1. Note that these are three TDs constructed from the single data set discussed in this chapter. Do the solutions agree within sampling errors? If they do not agree, what could explain the differences?

16.3 For the three TDs for $m = 6$,

ABD	ABC	ABE
ACF	ABF	ACD
ADE	AEF	ADF
BCE	BCD	BCF
BEF	CDE	BDE
CDF	DEF	CEF

conduct the analysis described in Example 16.1 using the barley grain weight given in Table 12.2 and barley plant weight data given in Table 15.9.

16.4 In Example 16.2, biomass data for the seven cultivars were used. Use only the biomass data for the six cultivars given in Table 16.5 and perform the calculations described for Example 16.2.

16.5 Suppose that the treatment design of Example 16.2 was used and that the mixture means were

$$y_{\cdot abdg} = 8, \quad y_{\cdot acfg} = 8, \quad y_{\cdot adeg} = 9, \quad y_{\cdot bceg} = 8, \quad y_{\cdot befg} = 9,$$

$$y_{\cdot cdfg} = 9, \quad y_{\cdot abcdefg} = 52/7.$$

Obtain the solutions for cultivar means and for $\mu+\delta.$ as described in Example 16.2.

16.6 For the barley weight data in Table 16.8, use the PBIBD treatment designs in Problem 16.2 and add the barley weight for the mixture with six cultivars and obtain the calculations decribed for Example 16.2.

16.7 Using the TD described in Example 16.3 and the barley grain weight from Table 12.2, conduct the calculations described in Example 16.3. Do likewise for biomass of the six cultivars and for all seven.

16.8 Barley plant weight was used in Example 16.4. Use the barley grain weight data from Table 12.2 and perform the calculations described for Example 16.4.

16.9 Add the following data to that of Table 16.7 and perform the analysis described for Example 16.4:

Mixture	Replicate 1	Replicate 2	Replicate 3
ABC	52.98	68.56	56.17
ABF	48.81	48.01	48.98
ADE	44.05	54.40	58.32
BCD	62.18	59.87	47.95
CEF	70.08	58.54	66.83

16.9 Literature Cited

1. Balaam, L.N. (1986). Intercropping—past and present. In *Statistical Design. Theory and Practice, Proceedings of Conference in Honor of Walter T. Federer* (Editors: C. E. McCulloch, S. J. Schwager, G. Casella, and S. R. Searle), Biometrics Unit, Cornell University, Ithaca, NY, pp. 141-159.
2. Bush, K. A., W. T. Federer, H. Pesotan, and D. Raghavarao (1984). New combinatorial designs and their application to group testing. *J. Statist. Plan. Inference* **10**, 335-343.
3. Denes, J. and A. D. Keedwell (1974). *Latin Squares and Their Applications*. Halstead Press, New York.
4. Federer, W. T. (1967). *Experimental Design–Theory and Application*. Oxford & IBH Publishing Co., Bombay. (First published by Macmillan, 1955.)
5. Federer, W. T. (1979). Statistical design and analysis for mixtures of cultivars. *Agron. J.* **71**, 701-706.
6. Federer, W. T. (1991). *Statistics and Society—Data Collection and Interpretation*. Marcel Dekker, Inc., New York, Chapter 5.
7. Federer, W. T. (1995). A simple procedure for constructing experiment designs with incomplete blocks of sizes 2 and 3. *Biometr. J.* **37**, 899-907.
8. Federer, W. T. and K. E. Basford (1991). Competing effects designs and models for two-dimensional layouts. *Biometrics* **47**, 1461-1472.
9. Federer, W. T., A. Hedayat, C. C. Lowe, and D. Raghavarao (1976). Application of statistical design theory to crop estimation with special reference to legumes and mixtures of cultivars. *Agron. J.* **68**, 914-919.
10. Federer, W. T. and D. Raghavarao (1987). Response models and minimal designs for mixtures of n of m items useful for intercropping and other investigations. *Biometrika* **74**, 571-577.
11. Fisher, R. A. and F. Yates (1938). *Statistical Tables for Biological, Agricultural and Medical Research*. Oliver and Boyd Ltd, Edinburgh (5th edn. 1957.)
12. Hall, D. A. (1976). Mixing design: A general model, simplifications, and some minimal designs. M.S. Thesis, Cornell University, Ithaca, NY
13. Hedayat, A. and W. T. Federer (1970). On the equivalence of Mann's group automorphism method of constructing $0(n, n-1)$ set and Raktoe's collineation method of constructing a balanced set of ℓ-restrictional prime-powered lattice designs. *Ann. Math. Statist.* **41**, 1530-1540.
14. Hedayat, A. and W. T. Federer (1974). Pairwise and variance balanced incomplete block designs. *Ann. Inst. Statist. Math.* **26**, 331-338.
15. Hedayat, A. and W. T. Federer (1984). Orthogonal F-rectangles for all even n. *Calcutta Statist. Assoc. Bull.* **33**, 85-92.
16. Jensen, N. F. and W. T. Federer (1965). Competing ability in wheat. *Crop Sci.* **5**, 449-452.
17. Khare, M. and W. T. Federer (1981). A simple construction procedure for resolvable incomplete block designs for any number of treatments. *Biometr. J.* **23**, 121-132.

18. Martin, K. J. (1980). A partition of a two-factor interaction, with an agricultural example. *Appl. Statist.* **29**, 149-155.

19. Raghavarao, D. (1971). *Constructions and Combinatorial Problems in Design of Experiments*. John Wiley & Sons, Inc., New York.

20. Raghavarao, D. and W. T. Federer (1979). Block total response as an alternative to the randomized response method in surveys. *J. Roy. Statist. Soc., Series B* **41**, 40-45.

21. Raghavarao, D., W. T. Federer, and S. J. Schwager (1986). Characteristics distinguishing among balanced incomplete block designs with respect to repeated blocks. *J. Statist. Plan. Inference* **13**, 151-163.

22. Raghavarao, D. and J.B. Wiley (1986). Testing competing effects among soft drink brands. In *Statistical Design. Theory and Practice, Proceedings of Conference in Honor of Walter T. Federer* (Editors: C. E. McCulloch, S. J. Schwager, G. Casella, and S. R. Searle). Biometrics Unit, Cornell University, Ithaca, NY, pp. 161-176.

23. Searle, S. R. (1979). On inverting circulant matrices. *Linear Alg. Applic.* **25**, 77-89.

24. Smith, L. L., W. T. Federer, and D. Raghavarao (1975). A comparison of three techniques for eliciting truthful answers to sensitive questions. *Proceedings, American Statistical Association, Social Statistics Section*, pp. 447-452.

25. Snedecor, G. W. and W. G. Cochran (1980). *Statistical Methods*, 7th edn., Iowa State University Press, Ames, IA.

Appendix 16.1

Federer and Raghavarao (1987) considered the class of BIBDs which was constructed by taking all possible combinations of m items taken n at a time, or the number of combinations was $v = m!/n!(m-n)!$. For each n, they considered that main effects and all interactions up to an n-factor, or $(n-1)$th-order, interaction were to be estimated. They did not consider the class of minimal TDs for estimating cultivar (main) effects and all interactions up to a k-factor interaction or mixing effect for $k < n$. The type of balance required in order to obtain relatively simple analytic results in terms of means or $n > k$ is that pairwise balance, λ_2, is required for cultivar effects; λ_2 plus triplet balance, λ_3, is required for cultivar plus BSMA effects; λ_2 plus λ_3 plus quartet balance, λ_4, is required for cultivar plus BSMA plus TSMA effects; etc. For $n = k$, one higher balance requirement may be omitted. Experiment design theory for t designs (Bush *et al.*, 1984) may be used to construct minimal balanced designs up to the order of balance required. The results given below are for the above type of balanced TDs. The restrictions on the solutions used are that the sum of the cultivar effects is zero, the sum of BSMA effects for cultivar h is zero, the sum of the TSMA effects for any pair hi is zero, etc. A computational procedure for using these restrictions for BSMA and TSMA effects is given in Appendix 16.2.

When the TD is not of the balance described above, analytic solutions are more complicated. The numerical results may be obtained using software packages such as GAUSS. This was demonstrated in Examples 16.3 and 16.4. If analytic solutions are desired, it is suggested that use be made of such packages as MAPLE and MATHEMATICA, as was done in Chapter 15. The tedious algebraic manipulations are thus alleviated.

$n = 2$: TD = BIBD with $v = m(m-1)/2$ and $\lambda = 1$, Cultivar + BSMA Effects

The solution for cultivar and BSMA effects when the TD is a balanced incomplete block design with m cultivars in all combinations of two results in $v = m(m-1)/2$ mixtures. For this TD, $\lambda = 1$ and each cultivar occurs in $s = m - 1$ mixtures. Using the usual restrictions that the sum of the effects is zero and response equation (16.6), the equations for the mean are

$$\bar{y}_{\cdot hi} = \mu + (\gamma_h + \gamma_i)/2 + \pi_{hi} + \sum_{g=1}^{r} \epsilon_{ghi}/r,$$

$$\bar{y}_{\cdot h \cdot} = \mu + (m-2)\,\gamma_h/2(m-1) + \sum_{g=1}^{r} \sum_{\alpha, h\epsilon S_\alpha}^{m} \epsilon_{ghi}/r(m-1),$$

$$\bar{y}_{\cdot\cdot i} = \mu + (m-2)\,\gamma_i/2(m-1) + \sum_{g=1}^{r} \sum_{\alpha, i\epsilon S_\alpha}^{m} \epsilon_{ghi}/r(m-1),$$

$$\bar{y}_{\cdots} = \mu + 2 \sum_{g=1}^{r} \sum_{\alpha,h,i \epsilon S_\alpha}^{m} \epsilon_{ghi}/rm(m-1).$$

Note that μ above is equal to $\mu + \delta$ in the preceding text and that the cultivar effect γ_h is equal to $\tau_h + \delta_h - \delta.$ in the preceding text. Also, we use the effect over n, the mixture size, whereas Federer and Raghavarao (1987) do not. The solutions obtained from the above are

$$\hat{\gamma}_h = \frac{2(m-1)(\bar{y}_{\cdot h\cdot} - \bar{y}_{\cdots})}{(m-2)}$$

and

$$\hat{\pi}_{hi} = \bar{y}_{\cdot hi} + \frac{m}{m-2}\,\bar{y}_{\cdots} - \frac{m-1}{m-2}\,(\bar{y}_{\cdot h\cdot} + \bar{y}_{\cdot i\cdot}).$$

The variance of a difference between two cultivar means or effects h and h' is

$$\begin{aligned}\mathrm{Var}(\hat{\gamma}_h - \hat{\gamma}_{h'}) &= \mathrm{Var}\left(\frac{2(m-1)}{m-2}\,(\bar{y}_{\cdot h\cdot} - \bar{y}_{\cdot i\cdot}\right)\\ &= \frac{4(m-1)^2}{(m-2)^2}\,\sigma_\epsilon^2\,\frac{2(m-2)}{r^2(m-1)^2} = \frac{8}{r(m-2)}\,\sigma_\epsilon^2.\end{aligned}$$

The variance of a difference between two BSMA effects hi and hi' is

$$\mathrm{Var}(\hat{\pi}_{hi} - \hat{\pi}_{hi'}) = \mathrm{Var}\left(\bar{y}_{\cdot hi} - \bar{y}_{hi'} - \frac{m-1}{m-2}\,(\bar{y}_{\cdot i\cdot} - \bar{y}_{\cdot i'\cdot})\right) = \frac{2(m-3)\sigma_\epsilon^2}{r(m-2)}.$$

$n = 3$: *TD = BIBD with* $v = m(m-1)/2$ *Mixtures, Cultivar + BSMA Effects*

For mixtures of size $n = 3$ and a BIBD TD, the number of mixtures v is equal to $m(m-1)/2$ mixtures, $s = 3(m-1)/2$ (m odd) occurrences of a cultivar in mixtures, and $\lambda = 3$ occurrences of pairs of cultivars in mixtures. The various means in terms of the parameters in response equation (16.7) are

$$\bar{y}_{\cdot hij} = \mu + \frac{(\gamma_h + \gamma_i + \gamma_j)}{3} + \frac{2(\pi_{hi} + \pi_{hi} + \pi_{hi})}{3} + \sum_{g=1}^{r} \frac{\epsilon_{ghij}}{r},$$

$$\bar{y}_{\cdot hi\cdot} = \mu + \frac{m-3}{3(m-2)}(\gamma_h + \gamma_i) + \frac{2(m-4)}{3(m-2)}\pi_{hi} + \sum_{g=1}^{r} \sum_{\alpha,hi \epsilon S_\alpha} \frac{\epsilon_{ghij}}{(m-2)},$$

$$\bar{y}_{\cdot h\cdot j} = \mu + \frac{m-3}{3(m-2)}(\gamma_h + \gamma_j) + \frac{2(m-4)}{3(m-2)}\pi_{hj} + \sum_{g=1}^{r} \sum_{\alpha,hj \epsilon S_\alpha} \frac{\epsilon_{ghij}}{(m-2)},$$

$$\bar{y}_{\cdot\cdot ij} = \mu + \frac{m-3}{3(m-2)}(\gamma_i + \gamma_j) + \frac{2(m-4)}{3(m-2)}\pi_{ij} + \sum_{g=1}^{r} \sum_{\alpha,ij \epsilon S_\alpha} \frac{\epsilon_{ghij}}{(m-2)},$$

$$\bar{y}_{\cdot h\cdot\cdot} = \mu + \frac{m-3}{3(m-1)}\,\gamma_h + 2\sum_{g=1}^{r}\sum_{\alpha,h\epsilon S_\alpha} \frac{\epsilon_{ghij}}{(m-1)(m-2)},$$

$$\bar{y}_{\cdot\cdot i\cdot} = \mu + \frac{m-3}{3(m-1)}\,\gamma_i + 2\sum_{g=1}^{r}\sum_{\alpha,i\epsilon S_\alpha} \frac{\epsilon_{ghij}}{(m-1)(m-2)},$$

$$\bar{y}_{\cdots j} = \mu + \frac{m-3}{3(m-1)}\,\gamma_j + 2\sum_{g=1,j\epsilon S_\alpha}^{r} \frac{\epsilon_{ghij}}{(m-1)(m-2)},$$

$$\bar{y}_{\cdots\cdot} = \mu + 6\sum_{g=1}^{r}\sum\sum_{h<i<j}\sum \frac{\epsilon_{ghij}}{m(m-1)(m-2)}.$$

Solutions for $\hat{\gamma}_h$ and $\hat{\pi}_{hi}$ are

$$\hat{\gamma}_h = \frac{3(m-1)}{(m-3)}\,(\bar{y}_{\cdot h\cdot\cdot} - \bar{y}_{\cdots\cdot})$$

and

$$\hat{\pi}_{hi} = \frac{3(m-2)}{2(m-4)}\bar{y}_{\cdot hi\cdot} - \frac{3(m-1)}{2(m-4}\,(\bar{y}_{\cdot h\cdot\cdot} + \bar{y}_{\cdot\cdot i\cdot}) + \frac{3m}{2(m-4)}\,\bar{y}_{\cdots\cdot\cdot}.$$

$n = 3$: *TD = BIBD, GMA + BSMA + TSMA*, $v = m(m-1)(m-2)/6$

Using response equations (16.9) and the above restrictions, the equations for the various means are

$$\bar{y}_{\cdot hij} = \mu + \frac{(\gamma_h + \gamma_i + \gamma_j)}{3} + \frac{2(\pi_{hi} + \pi_{hj} + \pi_{ij})}{3} + \eta_{hij} + \sum_{g=1}^{r} \frac{\epsilon_{ghij}}{r},$$

$$\bar{y}_{\cdot hi\cdot} = \mu + \frac{(m-3)(\gamma_h + \gamma_i)}{3(m-2)} + \frac{2(m-4)\,\pi_{hi}}{3(m-2)} + \sum_{g=1}^{r}\sum_{\alpha;h,i\epsilon S_\alpha}^{m} \frac{\epsilon_{ghi}}{r(m-2)},$$

$$\bar{y}_{\cdot h\cdot j} = \mu + \frac{(m-3)(\gamma_h + \gamma_i)}{3(m-2)} + \frac{2(m-4)\,\pi_{hj}}{3(m-2)} + \sum_{g=1}^{r}\sum_{\alpha;h,j\epsilon S_\alpha}^{m} \frac{\epsilon_{ghi}}{r(m-2)},$$

$$\bar{y}_{\cdot\cdot ij} = \mu + \frac{(m-3)(\gamma_h + \gamma_i)}{3(m-2)} + \frac{2(m-4)\,\pi_{ij}}{3(m-2)} + \sum_{g=1}^{r}\sum_{\alpha;i,j\epsilon S_\alpha}^{m} \frac{\epsilon_{ghi}}{r(m-2)},$$

$$\bar{y}_{\cdot h\cdot\cdot} = \mu + \frac{(m-3)\gamma_h}{3(m-1)} + 2\sum_{g=1}^{r}\sum_{\alpha;h\epsilon S_\alpha}^{m} \frac{\epsilon_{ghij}}{r(m-1)(m-2)},$$

$$\bar{y}_{\cdot\cdot i\cdot} = \mu + \frac{(m-3)\gamma_i}{3(m-1)} + 2\sum_{g=1}^{r}\sum_{\alpha;i\epsilon S_\alpha}^{m} \frac{epsilon_{ghi}}{r(m-1)(m-2)},$$

$$\bar{y}_{\cdots j} = \mu + \frac{(m-3)\gamma_h}{3(m-1)} + 2\sum_{g=1}^{r}\sum_{\alpha;j\epsilon S_\alpha}^{m} \frac{\epsilon_{ghi}}{r(m-1)(m-2)},$$

$$\bar{y}_{\cdots\cdot} = \mu + 6\sum_{g=1}^{r}\sum_{\alpha;h,i,j\epsilon S_\alpha}^{m} \frac{\epsilon_{ghi}}{rm(m-1)(m-2)}.$$

Solutions for the cultivar, BSMA, and TSMA effects are those given by Federer and Raghavarao (1987) with the definition of effects as described in the preceding text. A solution for a cultivar effect is

$$\hat{\gamma}_h = 3(m-1)(\bar{y}_{\cdot h\cdot\cdot} - \bar{y}_{\cdots\cdot}/(m-3).$$

A solution for a BSMA effect is

$$\hat{\pi}_{hi} = \frac{3(m-2)}{2(m-4)}\left\{\bar{y}_{\cdot hi\cdot} - \frac{(m-1)}{(m-2)}(\bar{y}_{\cdot h\cdot\cdot} + \bar{y}_{\cdot\cdot i\cdot}) + \frac{m}{(m-2)}\bar{y}_{\cdots\cdot}\right\}.$$

Appendix 16.2

A GAUSS program for obtaining results for Example 16.4 is presented below. Comments or annotations to the program appear in parentheses. Note that what is written inside the parenthesis will appear on the output. This makes it easy to discern the output entries. "This is a program for Example 16.4"; "Y is a vector of n times the mixture mean. For this case, these are the mixture totals."; Let Y[20,1] = 68.50 31.97 33.05 20.90 62.76 72.54 78.06 30.03 24.46 42.26 37.46 89.94 59.08 15.79 2.53 11.92 62.32 71.89 63.17 11.68; "The design matrix X will be X1 concatenated with X2 as this is easier.";

```
Let X1[20,21] =
1 1 1 0 0 0   2 2 0 0 0   2 0 0 0   0 0 0   0 0 0
1 1 0 1 0 0   2 0 2 0 0   0 2 0 0   0 0 0   0 0 0
1 1 0 0 1 0   2 0 0 2 0   0 0 2 0   0 0 0   0 0 0
1 1 0 0 0 1   2 0 0 0 2   0 0 0 2   0 0 0   0 0 0
1 0 1 1 0 0   0 2 2 0 0   0 0 0 0   2 0 0   0 0 0
1 0 1 0 1 0   0 2 0 2 0   0 0 0 0   0 2 0   0 0 0
1 0 1 0 0 1   0 2 0 0 2   0 0 0 0   0 0 2   0 0 0
1 0 0 1 1 0   0 0 2 2 0   0 0 0 0   0 0 0   2 0 0
1 0 0 1 0 1   0 0 2 0 2   0 0 0 0   0 0 0   0 2 0
1 0 0 0 1 1   0 0 0 2 2   0 0 0 0   0 0 0   0 0 2
0 1 1 1 0 0   0 0 0 0 0   2 2 0 0   2 0 0   0 0 0
0 1 1 0 1 0   0 0 0 0 0   2 0 2 0   0 2 0   0 0 0
0 1 1 0 0 1   0 0 0 0 0   2 0 0 2   0 0 2   0 0 0
0 1 0 1 1 0   0 0 0 0 0   0 2 2 0   0 0 0   2 0 0
0 1 0 1 0 1   0 0 0 0 0   0 2 0 2   0 0 0   0 2 0
0 1 0 0 1 1   0 0 0 0 0   0 0 2 2   0 0 0   0 0 2
0 0 1 1 1 0   0 0 0 0 0   0 0 0 0   2 2 0   2 0 0
0 0 1 1 0 1   0 0 0 0 0   0 0 0 0   2 0 2   0 2 0
0 0 1 0 1 1   0 0 0 0 0   0 0 0 0   0 2 2   0 0 2
0 0 0 1 1 1   0 0 0 0 0   0 0 0 0   0 0 0   2 2 2; X2=eye(20); X=X1~(3*X2);
let R[41,1]=.1.1.1.1.1.1   .125.125.125.125.125.125.125.125.125.125.125.125
.125.125.125   .33333.33333.33333.33333.33333.33333.33333.33333.33333
.33333.33333.33333.33333.33333.33333.33333.33333.33333.33333.33333;
Z=zeros(41,41);format 2,6;
"The various means are:";X'*(Y/3)'*DIAGRV(Z,R);
"J1 is the restriction on the BSMA effects to sum to zero."; Let J1[16,15]=
1 1 1 1 1   0 0 0 0   0 0 0   0 0 0
1 0 0 0 0   1 1 1 1   0 0 0   0 0 0
0 1 0 0 0   1 0 0 0   1 1 1   0 0 0
0 0 1 0 0   0 1 0 0   1 0 0   1 1 0
0 0 0 1 0   0 0 1 0   0 1 0   1 0 1
0 0 0 0 1   0 0 0 1   0 0 1   0 1 1;
"J2 in the restriction on the TSMA effects to sum to zero."; Let J2[15,20]=
1 1 1 1   0 0 0   0 0   0      0 0 0   0 0   0   0 0 0 0
```

```
1 0 0 0  1 1 1  0 0  0    0 0 0  0 0  0  0 0 0 0
0 1 0 0  1 0 0  1 1  0    0 0 0  0 0  0  0 0 0 0
0 0 1 0  0 1 0  0 1  0    1 1 0  0 0  0  0 0 0 0
0 0 0 1  0 0 1  0 0  1    0 1 1  0 0  0  0 0 0 0
1 0 0 0  0 0 0  0 0  0    1 1 1  0 0  0  0 0 0 0
0 1 0 0  0 0 0  0 0  0    1 0 0  1 1  0  0 0 0 0
0 0 1 0  0 0 0  0 0  0    1 0 0  1 0  1  0 0 0 0
0 0 0 1  0 0 0  0 0  0    0 1 0  1 0  1  0 0 0 0
0 0 0 0  1 0 0  0 0  0    1 0 0  0 0  0  1 1 0 0
0 0 0 0  0 1 0  0 0  0    0 1 0  0 0  0  1 0 1 0
0 0 0 0  0 0 1  0 0  0    0 0 1  0 0  0  0 1 1 0
0 0 0 0  0 0 0  1 0  0    0 0 0  1 0  0  1 0 0 1
0 0 0 0  0 0 0  0 1  0    0 0 0  0 1  0  0 1 0 1
0 0 0 0  0 0 0  0 0  1    0 0 0  0 0  1  0 0 1 1;
J=zeros(41,6)~(J1~zeros(6,20))||(zeros(35,15)~(J2:zeros(20,20)));
var=inv(X'*X - J);
"b is the vector of solutions for cultivar means, BSMA effects, and TSMA
effects.";
b=var*X'*Y/r; b';
"Sum of means";X'*Y/r;
S=ones(1,20);
"Since Y is a vector of mixture totals, the grand total is";
S*Y;
"Mixture sum of squares is";
trss=Y'*Y/r – (S*Y)*S*Y/60; trss;
b'*X'*Y/r;
Y'*Y/r;
"Cultivar sum of squares for C1 a vector of sums of means and C2 of cultivar
effects";
Let C1[6,1]=154.8433 120.7133 218.9067 116.9633 141.2333 128.6500;
Let C2[6,1]=3.9792 - 13.0858 - 13.0858 36.0108 - 14.9608 - 2.8258 - 9.1175;
C1'*C2;
"BSMA sum of squares";
Let B1[15,1] = 102.9467 187.9067 99.4800 118.5867 110.4533 163.9867 58.5000
94.4667 62.9533 156.2867 185.9800 181.4667 79.8800 73.7067 86.0200;
Let B2[15,1] = 1.3695 -3.5930 1.4758 -0.4605 1.2083 -1.0930 3.2932
-3.8055 -1.2455 0.7882 3.8145 -0.7705 1.6332 -2.8505;
B1'*B2;
"TSMA sum of squares";
Let T1[20,1] = 0.5019 2.8225 -2.4958 -0.8286 0.1303 -0.7197 0.0875 -0.2392
-2.7136 3.4547 -3.4547 2.7136 0.2392 -0.0875 0.7197 -0.1303 0.8286 2.4958 -
2.8225 -0.5019;
Y'*T1;
```

Appendix 16.3

The following is a MAPLE program to aid in the solution of the effects in Section 16.7.1:

```
eqs1:={u=y1,u+(m-3)*g12/(2*(m-2))=y12,u+
(m-3)*g13/(2*(m-2))=y13,u+(g12+g13)
/2+p=y123};a1:=solve(eqs1,{u,g12,g13,p});h1:=
collect(a1,{y1,y12,y13,y123},factor );
```

The following is a MATHEMATICA program used for obtaining the variances for the effects given in Section 16.7.1, where m was set equal to 5, 6, and 7:

```
m=5;
s1=2*(Sum[e[1,i,j],{i,2,m-1},{j,i+1,m}])/((m-1)*(m-2))
s2=2*(Sum[e[1,2,j],{j,3,m}]+Sum[e[2,i,j],
{i,3,m-1},{j,i+1,m}])/((m-1)*(m-2))
s3=2*(e[1,2,3]+Sum[e[1,3,j],{j,4,m}]+Sum[e[2,3,j],{j,4,m}]+
Sum[e[3,i,j],{i,4,m-1},{j,i+1,m}])/((m-1)*(m-2)
s12=Sum[e[1,2,j],{j,3,m}]/(m-2))
s13=(e[1,2,3]+Sum[e[1,3,j],{j,4,m}])/(m-2)
s14=(Sum[e[1,i,4],{i,2,3}]+Sum[e[1,4,j],{j,5,m}])/(m-2)
res={e[h_,i_,j_] e[h_,i_,j_]->ER,e[h_,i_,j_] e[d_,f_,g_]->0};
x1=Simplify[s1-s2]
Expand[x1^2]/.res
x2=Simplify[((m-2)*(s12-s12))/(m-3)]
Expand[x2^2]/.res
x3=Simplify[e[1,2,3]-e[1,2,4]-(m-2)*(s13-s14)/(m-3)]
Expand[x3^2]/.res
```

To suppress output for a given statement, add a semicolon for MATHEMATICA and delete the semicolon for MAPLE.

CHAPTER 17

Spatial and Density Arrangements

17.1 Introduction

Spatial arrangement, density level, intimacy, and orientation of two crops in a mixture were discussed in Volume I. These aspects of forming an experimental unit (e.u.) to study the effects of these four factors is the subject of this chapter when the polyculture consists of n cultivars. The type and nature of the e.u. has many more ramifications than if only two cultivars make up the intercrop combination. The concepts and ideas used for mixtures of two cultivars are utilized and expanded for mixtures of n cultivars. The number of variations and complexity of arrangements increases with the number of cultivars in the mixture. In the next section, density per hectare is held constant while spatial arrangement and intimacy is varied. For $n = 3$, cultivars in a mixture 18 arrangements, 1 to 18, out of many are discussed. For $n = 4$, cultivars in a mixture, 6 arrangements, 19 to 24, are presented.

In Section 17.3, variation in density levels of each cultivar in the mixture is included along with spatial arrangement, intimacy, and orientation. First, a factorial arrangement for d_i density levels for each of the n cultivars is considered. Then, fractional replicates of the factorial are discussed. For n and/or d_i large, the number of combinations becomes large and unrealistic from a practical standpoint for many situations. A class of plans which are parsimonious with respect to experimental area, number of e.u.s, and material are discussed. These parsimonious arrangements or plans for an e.u. cover a wide range of spatial arrangements, densities, and intimacies of the cultivars in the mixture. Plans are also included which allow measurements of orientation effects as well. Statistical analyses for the various designs are presented in Section 17.4. These are a straightforward extension of those in Volume I. Some comments and discussion are presented in Section 17.5. A set of problems for the reader's understanding of the ideas and concepts of this chapter is presented in Section 17.6.

The emphasis on modeling yield–density relationships as discussed in Chapter 14 is not the focus of this chapter. Rather the emphasis is on constructing plans, arrangements, or designs to use in experimentation which then may be used to construct models for yield–density and other relationsips. However, the arrangements described can be of considerable value in determining the nature and form of density, spatial arrangement, intimacy, and orientation relationships with a response such as yield, crop value, total dry matter, caloric content, etc.

Several considerations are involved in the selection of density levels and row spacings for each of the n cultivars in a mixture. The number of levels and the range of levels of a factor for a cultivar are important. The range of density levels, say, should include levels lower and higher than would be useful in practice. This is because the endpoints of a regression function have considerable influence on the form and accuracy of a yield–density relation. A yield–density relation necessarily goes through the origin and, hence, *if it is known* that the relationship is linear, then an optimal design is to take all observations at the highest density possible. Since yield–density relations are expected to be unknown and nonlinear, it is necessary to have several levels of density. The points of inflexion for the nonlinear relation are unknown and, hence, the best procedure is to take several equally spaced values of density. Mead and Riley (1981) suggest using four points but that should depend on how much knowledge the experimenter has on yield–density relations in polyculture. Usually, there is no knowledge and therefore it is suggested that at least double their number be used as linear segments and plateaus as well as points of inflexion are useful and necessary information about the relationship. Using the plans suggested, it is easy to obtain many levels of a factor and still remain experimentally practical.

Haizel (1974) states that there are three methods for combining plant populations of the n cultivars in a mixture. The *additive method* involves using the sole crop populations for each of the n cultivars. In the *substitutive method*, the total plant population of an area planted to the mixture is the same as the same area planted to sole crops. The third method, *replacement series* (see Chapter 18), requires that a certain number of plants of one cultivar is equivalent to replacing one plant by another one. Criticisms (Kass, 1978) have been raised about all three methods, in that none of them will arrive at the optimum population levels for a mixture of n cultivars. Levels obtained by each of the three methods most likely should be included in the levels selected for a study of yield–density relationships for the n cultivars in the mixture.

With respect to spatial arrangement, an experimenter oftens selects the arrangement which is experimentally easiest to handle. This may be a mistake, in that the experimental conditions may not apply to what is done in practice. For example, the experimenter may grow the n cultivars in adjacent rows, whereas the agriculturist grows them all mixed together. The experimenter may grow beans and maize in separate rows, whereas the farmer grows them in the same row. Not only the mixtures used but their spatial arrangements as used by the farmer must be taken into consideration in setting up polyculture experiments (Kass, 1978).

17.2 Spatial Arrangements—Density per Hectare Constant

Spatial and density arrangements for *n* cultivars in a mixture can take on many and varied forms. Several types of arrangements for mixtures of two cultivars are described in Volume I. With three or more cultivars in a mixture, many different and varied forms of arrangements are possible. For mixtures of three cultivars and with density per hectare kept constant, a number of possible spatial arrangements are described below. Note that the density within a row may vary, but the density per unit area is kept constant for the following 24 arrangements. Suppose, for example, that the three crops are cassava, C, maize, M, and bean, B. Let the cassava rows be 1 m apart for the following spatial arrangements:

Arrangement 1

C B M B C B M B C B M

The maize and cassava rows are 1 m apart and there is a 0.5-m distance between each crop row.

Arrangement 2

C M B M C M B M C M B

The maize rows are 0.5 m apart and the cassava and bean rows are 1 m apart.

Arrangement 3

B M M B C B M M B C B M M

The distance between the bean and cassava rows is 0.125 m, the distance between the bean row and a maize row is 0.25 m, and the distance between the pair of maize rows between two cassava rows is 0.25 m.

Arrangement 4

CB M CB M CB M

CB means that cassava and bean are planted together in the same row. The distance between CB rows is 1 m, as is the distance between maize rows. The M and CB rows are 0.5 m apart.

Arrangement 5

C MB C MB C MB

MB means that maize and bean are planted together in the same row. The distance between MB rows is 1 m, as is the distance between cassava rows. The C and MB rows are 0.5 m apart.

Arrangement 6

CB MB CB MB CB MB

The distance between the CB rows is 1 m and the distance between MB rows is 1 m. The CB and MB rows are 0.5 m apart.

In the following arrangements 7–18, the cassava rows are 2 m apart, but the density in a row is twice that for arrangements 1–6.

Arrangement 7

C B M B M B C B M B M B C

The distance between maize rows is 1 m. The distance between C and B rows is 0.5 m. The distance from B to M and from M to B is 0.25 m, while, for B located in the middle of the two maize rows, the distance is 0.5 m from each maize row.

Arrangement 8

C B M B B B M B C B M

The rows are all 0.25 m apart.

Arrangement 9

C B B M M B B C B B

Except for the two maize rows which are 0.5 m apart, the remaining rows are 0.25 m apart.

Arrangement 10

C B M B M B M B C B M

Here, the rows are all 0.25 m apart with a distance of 0.5 m between M rows and between B rows.

Arrangement 11

C B B M B B C B B M

The distance from C to B and B to B is 0.25 m. The distance from B to M is 0.5 m and the distance between two maize and two cassava rows is 2 m.

Arrangement 12

C MB C MB C MB

All rows are 1 m apart.

Arrangement 13

CB M CB M CB M

All rows are 1 m apart.

Arrangement 14

CB MB CB MB CB MB

All rows are 1 m apart.

Arrangement 15

CB MB MB CB MB MB C

Here, the distance between CB and MB is 0.75 m and between the two MBs is 0.5 m.

Arrangement 16

C B C B M B M B C B C

All rows are 0.5 m apart, leaving the pair of C rows and the pair of M rows each 1 m apart, but with 2 m between pairs.

Arrangement 17

CB CB MB MB CB CB

All rows are 1 m apart with the pairs of rows 2 m apart.

Arrangement 18

CB CB CB CB MB MB CB CB CB CB CB MB

Here, the ratio of CB rows to MB rows is 4:2 and the rows are 1 m apart.

An arrangement like 18 has been found useful for growing cowpea, C, in Nigeria. Interspersing rows of cowpea with rows of soybean is effective in controlling insect damage on cowpea. For a maize, soybean (S), and cowpea mixture, an arrangement of the following nature is useful:

Arrangement 19

SM SM C C C C C C C SM SM C

where SM means soybean and maize planted in the same row with 0.5 (or 1) m between the SM rows and the other rows are 0.5 m apart.

Similar arrangements may be constructed for four cultivars in the mixture. Suppose the four crops are cassava (C), bean (B), maize (M), and melon (E). Some arrangements are given below.

Arrangement 20

CE CE B M M B CE CE B

The distance between CE and CE is 1 m, as is the distance between M and M. B is 0.5 m from CE and from M. C and E are planted in the same row. M and B may involve two crops per year.

Arrangement 21

C E C E M B M E C E C E

All rows are equally spaced at 0.5 m intervals.

Arrangement 22

C E C E MB B MB E C E C E

All rows are equally spaced at 0.5 m intervals. M and B are planted in the same row.

Arrangement 23

C B B C E M E M E C B B C

The C rows are 1 m apart, the B rows are 0.25 m apart, M and E are 0.5 m apart, and C and B are 0.375 m apart.

Arrangement 24

CE CE MB MB CE CE

All rows are 1 m apart, with C and E being planted in the same row and the same for M and B.

Note that the mixture need not be planted in rows, but zero or more crops may be in rows and the other crops broadcast over the area. Such is the case with orchards, where the trees are in rows and the mixture of crops under the trees is drilled in or broadcast over the area. Pasture mixtures are not planted in rows. Obtaining individual cultivar yields in a broadcast hay crop is tedious and labor-intensive, as hand separation of the cultivars is necessary. Planting in rows may make it easy to obtain yields from each crop in the mixture. However, this may not be what is done in practice.

Several competition designs of a balanced nature are discussed in Federer and Basford (1991) and in Volume I. Cultivars may occur as individual plants or as groups of plants such as hills of three or more maize plants in each hill. These designs are constructed to equalize occurrences of cultivars adjacent to each other. The designs may be constructed for any size of mixture and are useful in studying various association, competition, and mixing ability properties of mixtures. The reader is referred to Volume I for a discussion of the designs and their analyses.

In constructing and using any design or arrangement, it is necessary to consider the following:

(i) the spatial arrangement for each crop in the mixture,
(ii) the intimacy of pairs of crops in the mixture
(iii) density of each crop in the mixture, and perhaps
(iv) orientation with respect to the Sun.

These concepts are discussed by Mead and Riley (1981) and in Volume I for mixtures of two cultivars. The ideas are directly extendable to mixtures of three or more cultivars. *Intimacy* refers to closeness of crops in the experimental unit. Crops that are planted in the same row are 100% intimate, and those planted in rows far enough apart to not affect each other have zero intimacy. In arrangement 24, for example, C and E and M and B have 100% intimacy and those planted in rows far enough apart to not affect each other have zero intimacy. In arrangement 24, for example, C and E and M and B have 100% intimacy, whereas there is less intimacy between C and M, C and B, and M and E. In arrangement 19, S and M are 100% intimate, whereas the center rows of C have zero intimacy with S and M. The spatial arrangement of a crop varies throughout the above 24 arrangements. The density within rows was varied to maintain a constant density per hectare in the above 24 arrangements. With respect to *orientation*, which is the placement of rows with respect to the Sun, light-sensitive crops in a mixture may perform better if the rows are in a north-south direction. Such orientation may allow more sunlight to the leaves, especially if tall crops are mixed with short crops. Shade requiring crops may perform better for rows in an east-west direction. Disease incidence may be affected by row orientation in that north-south rows may have less humidity and more light, which may be detrimental to disease buildup.

17.3 Spatial Arrangements—Density Variable

In this section, crop arrangements with variable densities per hectare are considered. Selection of the range of densities for each crop requires considerable thought, as does the determination of which crop has varying densities. The density of one or more of the crops may be held constant while varying the densities of the remaining crops in the mixture. A treatment design that usually would be considered first is to determine density levels d_{ij} for $j = 1, \ldots, d_i$, for each crop i, and then use a factorial arrangement of $N = \Pi_{i=1}^{n} d_i$ density–crop combinations. For example, suppose that four density levels of cassava, three density levels of bean, and five density levels of maize were under consideration. The number of combinations is $4 \times 3 \times 5 = 60$. As is obvious, the number of combinations in a complete factorial becomes large quickly as n and d_i increase. Furthermore, yield–density relationships are ill-determined with small numbers of density levels, and increasing their number would greatly increase the number of combinations in the treatment design.

In place of using a complete factorial, a fractional replicate of a complete factorial could be considered. Saturated fractions are the most parsimonious and may be obtained for any resolution required. They are, however, not always variance-optimal, as a variance-optimal fraction usually needs to be constructed as plans are available for only a few situations. A saturated fraction, as many effect degrees of freedom as observations, is easily constructed using the one-at-a-time method described by Anderson and Federer (1975). For the above example and for a Resolution V fraction (all main effects and all two-factor interactions are estimable if all higher-order interactions do not exist), the treatment design for the order of crops as cassava, bean, and maize is

1 1 1	2 2 1	3 1 2	1 2 4
2 1 1	2 3 1	3 1 3	1 2 5
3 1 1	3 2 1	3 1 4	1 3 2
4 1 1	3 3 1	3 1 5	1 3 3
1 2 1	4 2 1	4 1 2	1 3 4
1 3 1	4 3 1	4 1 3	1 3 5
1 1 2	2 1 2	4 1 4	
1 1 3	2 1 3	4 1 5	
1 1 4	2 1 4	1 2 2	
1 1 5	2 1 5	1 2 3	

These 36 combinations of density levels of the 3 crops allow estimation of all main effects and all two-factor interactions. The three-factor interaction is assumed nonexistent. There are 3 + 2 + 4 = 9 degrees of freedom for main effects and 6 + 12 + 8 = 26 degrees of freedom for the 3 two-factor interactions. Thus, 9 + 26 = 35 plus one degree of freedom for the overall mean equals 36. As Anderson and Federer (1975) demonstrate, a one-at-a-time plan is least optimal for main effect plans or Resolution III plans. The variance optimality improves as the resolution of the fraction increases. The above fraction would cut the number of combinations from 60 to 36, i.e., approximately one-half but still relatively large.

If only main effects of say the four factors orientation, intimacy, spacing, and density, e.g., were of interest and if three levels, 0, 1, and 2, of each factor were being considered, the following nine factorial combinations forms an optimal saturated main effect, Resolution III, plan:

0000 0111 0222 1021 1102 1210 2012 2120 2201

For 5 factors at 4 levels, 0, 1, 2, and 3, of each factor, the following 16 factorial combinations form an optimal saturated main effect, Resolution III, design:

00000	01111	02222	03333	10132	11023	12310	13201
20213	21302	22031	23120	30321	31230	32103	33012

A much more parsimonious approach for studying density levels and spatial arrangements as they affect yield is available. Expanding upon the ideas of Nelder (1962), Federer and Scully (1992), and those in Volume I, many parsimonious plans or arrangements may be constructed. Within a *single* e.u., it is possible to include many density levels and various diverse spatial arrangements. Parsimony of experimental area and plant material is achieved at the expense of additional labor per e.u. Nelder's (1962) fan designs were called systematic plans in that the variation in density levels and/or spatial arrangements was systematic within an e.u. His designs were constructed for sole crop studies. The parsimonious experiment designs (PED) of Federer and Scully (1992) are of the same nature as Nelder's. Their idea was that PEDs would limit the number of sites for varietal trials by including as many sources of site-to-site variation as possible at a single site. Identifying site-to-site variation and including these in an experiment at a single site allows the experimenter to identify specific interaction components in a variety by site (genotype by environment) interaction. They identified such site-to-site sources of variation as biological time of planting, fertilizer level, density levels, disease levels, insect levels, water levels, etc. as candidates for study at a single site. Some sources of variation which usually cannot be included are temperature level, humidity level, length of growing season, and elevation. For sources of variation which can be studied at a single site, the performance of a cultivar with respect to the sources of variation included may be assessed. Without defining what is meant by site-to-site variation, a variety by site interaction is not interpretable and is meaningless for prediction of a variety's performance with respect to specific sources of variation.

A parsimonious plan for varying density levels and spatial arrangements for a mixture of three cultivars C, cassava, M, maize, and bean, B, is given in Figure 17.1. The distance between plants within a row of a cultivar is constant. The decrease in density per hectare is accompanied by an increase in distance between rows of the cultivars; i.e., their effects are completely confounded. By plotting plant yields against distance between rows or density per hectare, an estimate of the density–spacing combination resulting in highest yields may be obtained. Or, the distance beween plants within a row may be changed in such a manner as to keep the density per hectare a constant throughout the e.u. (see Section 8.3, Volume I). Also, another row of C parallel to the middle row C could be added and the plant distance be kept constant in the left part of the e.u. and varied in the right half in such a manner as to keep density per hectare constant. This would allow separation of density and spatial effects. Such arrangements allow yield–density and yield–spatial relationships to be evaluated and to select the density and spatial arrangement to maximize yield. The plan could be expanded to equalize the number of rows of C and M to have 2C:2M:4B rows in both parts of the e.u. An e.u. of the nature of the one in Figure 17.1 could be included in an experiment design, say a RCBD, where the treatments could be different lines of a cultivar, different fertilizer levels, different moisture levels, a subset of or all 24 arrangements of cultivars described above, etc. The optimal density–spacing combination for each treatment could be determined as the average over all replicates. The variation of

```
C        B        M        B       C   B  M   B  C
C       B        M       B       C   B    M  B   C
C       B       M       B      C    B    M   B   C
C      B       M      B      C    B     M   B    C
C      B      M      B      C    B     M    B    C
C     B      M     B      C    B      M    B     C
C     B     M     B     C     B      M     B     C
C    B     M    B     C      B      M     B      C
C    B    M    B    C      B       M      B      C
C   B    M   B    C       B       M       B      C
C   B   M   B   C       B        M       B       C
C  B   M  B   C        B        M        B       C
```

FIGURE 17.1. Rectangular experimental unit for varying spatial arrangements and density per hectare for three crops: cassava (C), bean (B), and maize (M). Distance between plants in a row is constant.

this response from e.u. to e.u. could also be obtained. It may be necessary to set up experimnts at a number of sites in order to obtain the optimal density–spacing combination recommendation for a region.

For certain situations where row distances are predetermined, only yield–density and not yield–spatial relationships are required. Arrangements such as those in Figure 17.2 may be used for this situation. This design was suggested by Federer and Scully (1992) and is in the spirit of the Nelder (1962) designs. Using any of the 24 or other arrangements in Section 17.2, the distance between plants decreases from lowest density to highest density down the row, but the distance between rows, spatial arrangement, remains constant. As in the Figure 17.1 e.u.s, this allows inclusion of all the densities to be studied in a single e.u., certainly a parsimonious situation with regard to land area. Although the total labor is considerably reduced, the labor per e.u. does increase as responses are required for each plant, or for groups of plants within an e.u., rather than a single response for an e.u. as required in factorial treatment designs. The planting and harvesting labor is less for the e.u. on the right in Figure 17.2. The one on the left has continuously increasing density down a row, whereas the one on the right has only l density levels increasing in a systematic manner. Even though the density was continuouly increasing down the row, the experimenter might harvest l segments of the row with the density level being the average density within the segment. If the stand is uneven, perhaps a covariate of distance between plants and/or number of plants per segment could be used to adjust the plant yields within each e.u. The adjusted yields would then be used for obtaining a yield–density or other relation.

Nelder (1962) introduced arrangements for an e.u. which are called fan designs for studying spatial–density–yield relationships. These designs are parsimonious relative to area and number of e.u.s in the manner of those discussed above. His arrangements are for single cultivars but are easily adapted to accommodate any

density increasing					density levels					
C	B	M	B	C	d_1	C	B	M	B	C
C	B	M	B	C		C	B	M	B	C
C	B	M	B	C		C	B	M	B	C
C	B	M	B	C	d_2	C	B	M	B	C
C	B	M	B	C		C	B	M	B	C
C	B	M	B	C		C	B	M	B	C
					....					
C	B	M	B	C	d_l	C	B	M	B	C
C	B	M	B	C		C	B	M	B	C
C	B	M	B	C		C	B	M	B	C

FIGURE 17.2. Two experimental units for increasing densities within rows of a cultivar for three cultivars C, B, and M. Density increases from low to high for the e.u. on the left and by l discrete levels for the e.u. on the right. Space between plants decreases continuously down the row for design on the left and by increments for the design on the right.

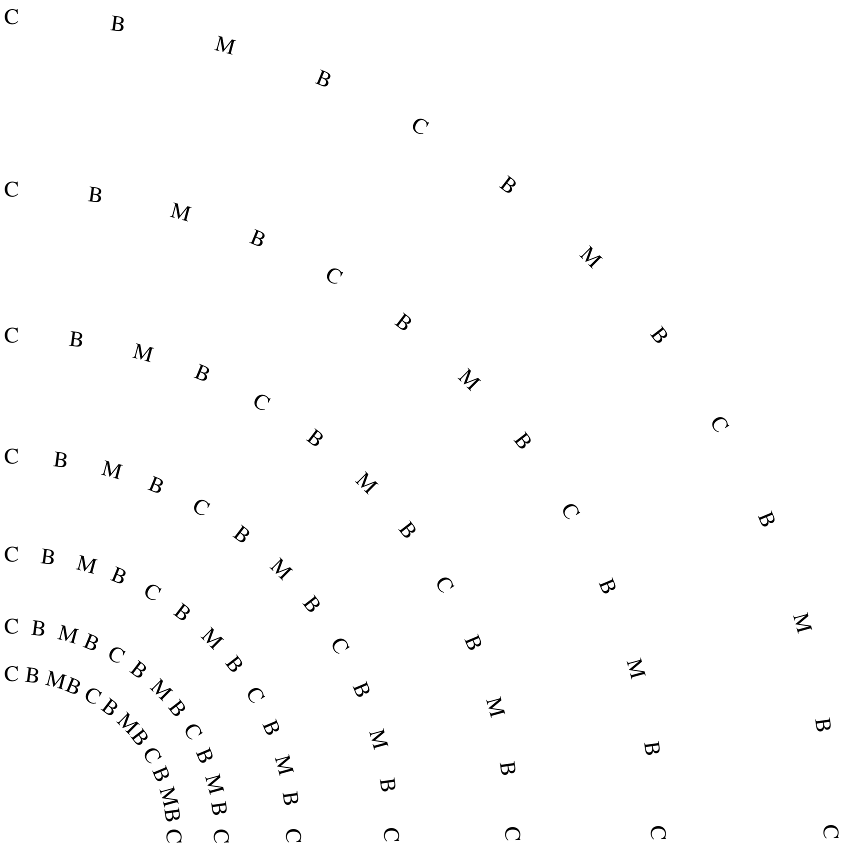

FIGURE 17.3. Fan design for increasing distance between plants within a row.

polyculture combination such as those described for arrangements 1–24. Some of his fan designs are adapted here for mixtures of cultivars. The rays from the origin in a quadrant form the rows of a cultivar in the mixture such as depicted in Figure 17.3. Here, a mixture of three cultivars with increasing distance between plants in a row is used to illustrate the procedure. Also, the distance between the rows is increasing. The rows form rays of a quadrant of a circle and the cultivars on any point on a ray equidistant from the origin, an arc, form the sequence in the crop arrangement. The distance from the origin forms the density. The highest density per hectare are those closest to the origin. Those farthest from the origin have the lowest densities per hectare. The intimacy of cultivars decreases as the distance from the origin increases, i.e., the distance between cultivars increases. Increasing space between rows and decreasing density per hectare are confounded for this fan design. The optimum response is determined for each row (cultivar) of the design in the e.u. The design may be varied by using equally spaced double rows for

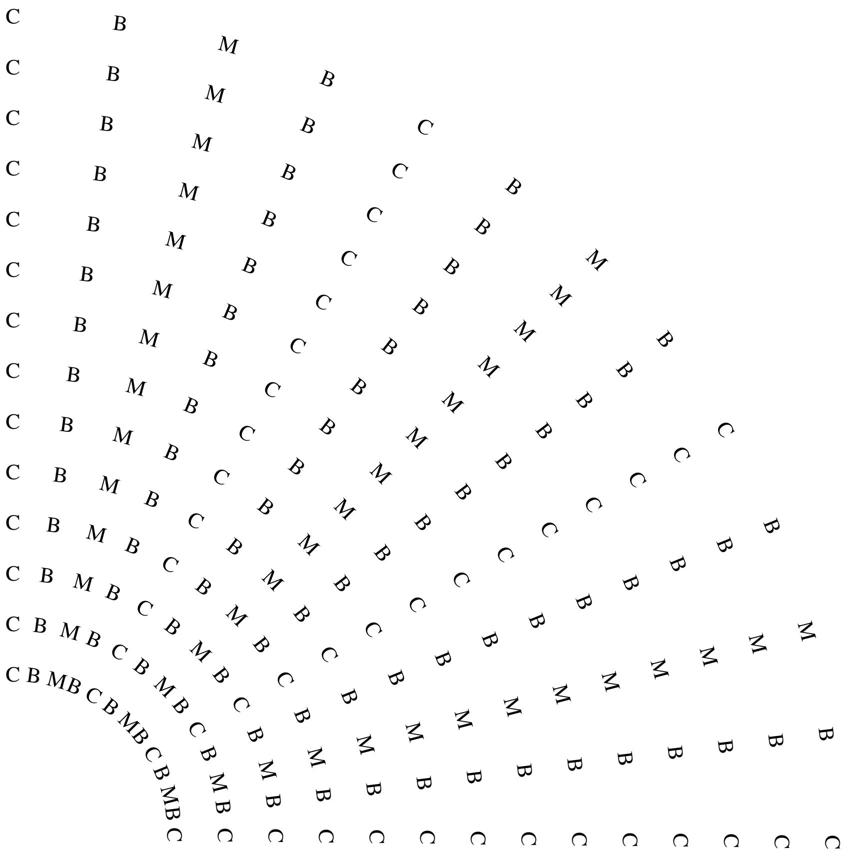

FIGURE 17.4. Fan design for plants of three cultivars equally spaced within each row.

some or all of the crops in the polyculture. With the additional plants per cultivar, a yield–density–spacing relation is better determined.

Instead of having the distance between plants within a row increase with distance from the origin, the distance may remain constant, as in Figure 17.4. The density per hectare does decrease but not as fast as in Figure 17.3. From a planting standpoint, the plan in Figure 17.4 is more desirable and less cumbersome than the one in Figure 17.3. The row spacings in the two plans are comparable. Alternatively, the distance between plants could decrease as distance from the origin increases in such a manner as to keep density per hectare constant on all arcs of the quadrant. Since density per hectare is constant, only spatial arrangement will vary. A method for determining circles with constant areas is described in Section 8.3, Volume I.

Two particular forms of the Nelder fan design adapted for polycultures have been presented. Another form has been given by Hiebsch *et al.* (1995) in a study of plant density and soybean maturity. From these, it may be noted that any mixture of n cultivars may be used in such a plan. One further note is that the area of the fan

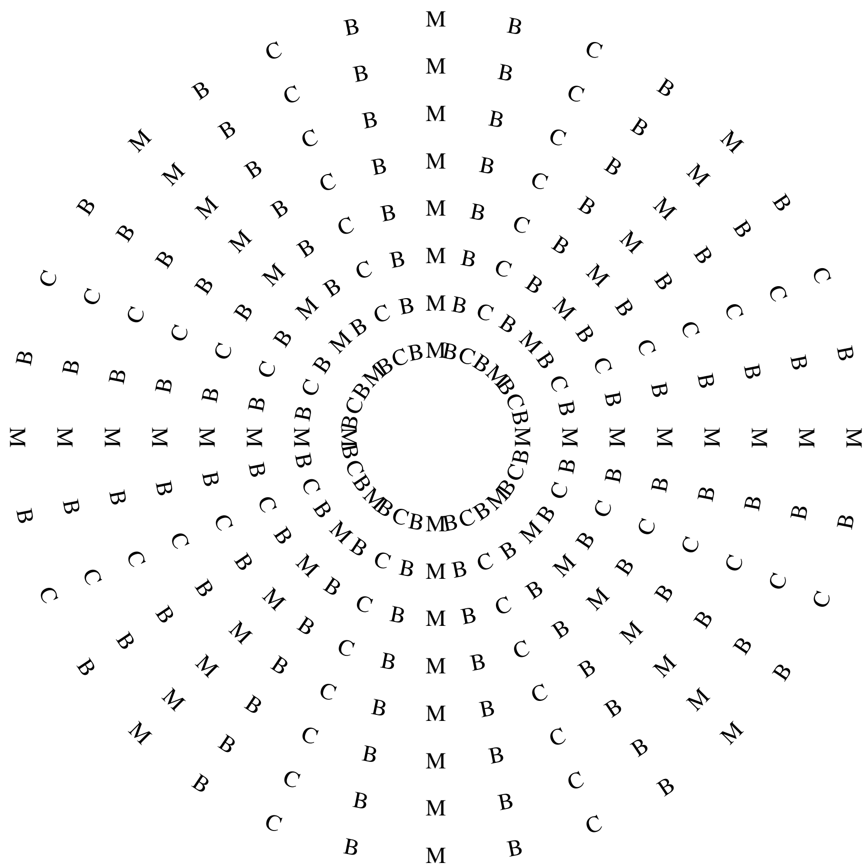

FIGURE 17.5. Okigbo circle design for three cultivars equally spaced within the row.

design is not rectangular, as might be required for field arrangements of e.u.s. It is easy to make a rectangular e.u. simply by extending the rows to form a rectangle in Figures 17.3 and 17.4. Alternatively, it may be necessary to eliminate the border effect by using a border around the e.u. The material used for the border could be used to form a rectangular area.

If the cultivars are all mixed together, i.e., 100% intimacy, the Nelder (1962) designs may be used directly and treated as a single cultivar in forming the plan.

In order to consider row orientation with respect to direction, the Okigbo (1978) circle designs discussed in Volume I may be extended to include mixtures of n cultivars. Here, we consider that individual rows of the cultivars will be planted. If all n cultivars are intermixed, then individual row arrangements need not be considered. If the top of Figure 17.5 is considered to be north, the maize rows would most likely be laid out as shown, as maize is somewhat light sensitive and the biggest contrast then would be between the north-south and east-west maize rows. The number of rows in a circle will depend on the various row spacings

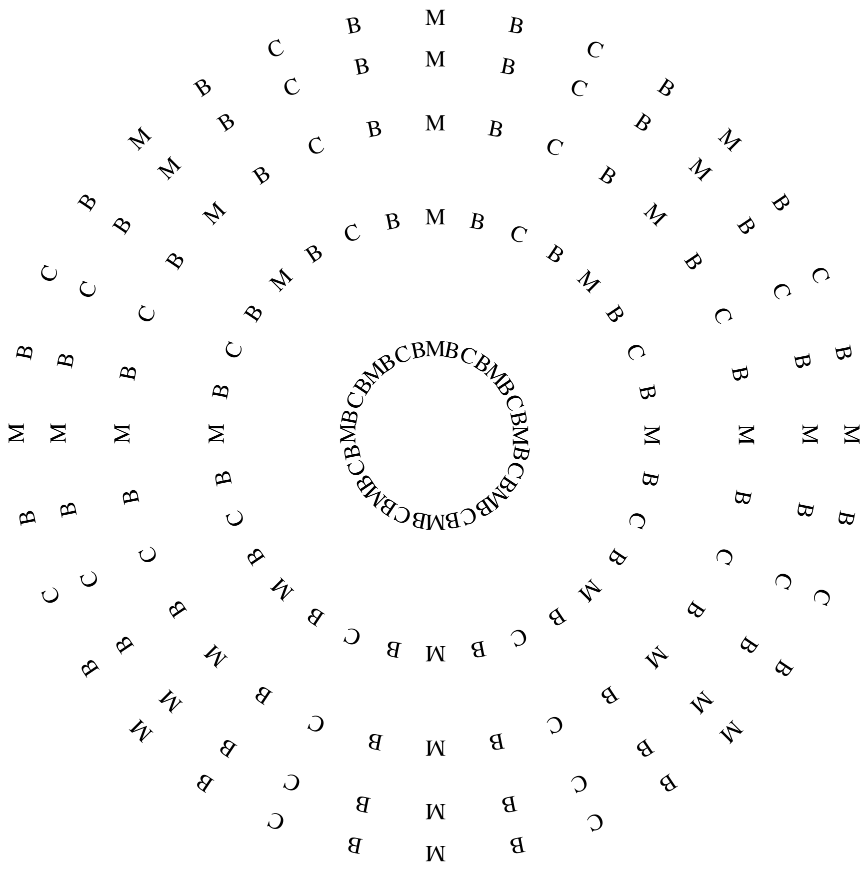

FIGURE 17.6. Okigbo circle design for three cultivars with decreasing distance between the plants within a row but increasing distance between rows, keeping the density per hectare constant.

under consideration. The radius of the inner circle where none of the mixture is planted also needs to be considered. It has to be large enough to allow a reasonable narrowest row spacing as a starting point and to allow planting and harvesting operations to function. With this plan, it is possible to have one plan in one-half of the circle and another plan in the other half. It would still be possible to study orientation effects. In Figure 17.5, the distance between plants within a row (ray) will be constant, whereas the distance between rows is increasing. The density per hectare of each cultivar is decreasing. As described above, a rectangular e.u. may be obtained by extending the rows to a rectangular border. The advantage of the circle over the fan design is that orientation effects in addition to row spacing–density effects are obtained.

In Figure 17.6, the distance between plants is decreased to keep the density per hectare constant, thus allowing only row spacing effects to be assessed. This plan

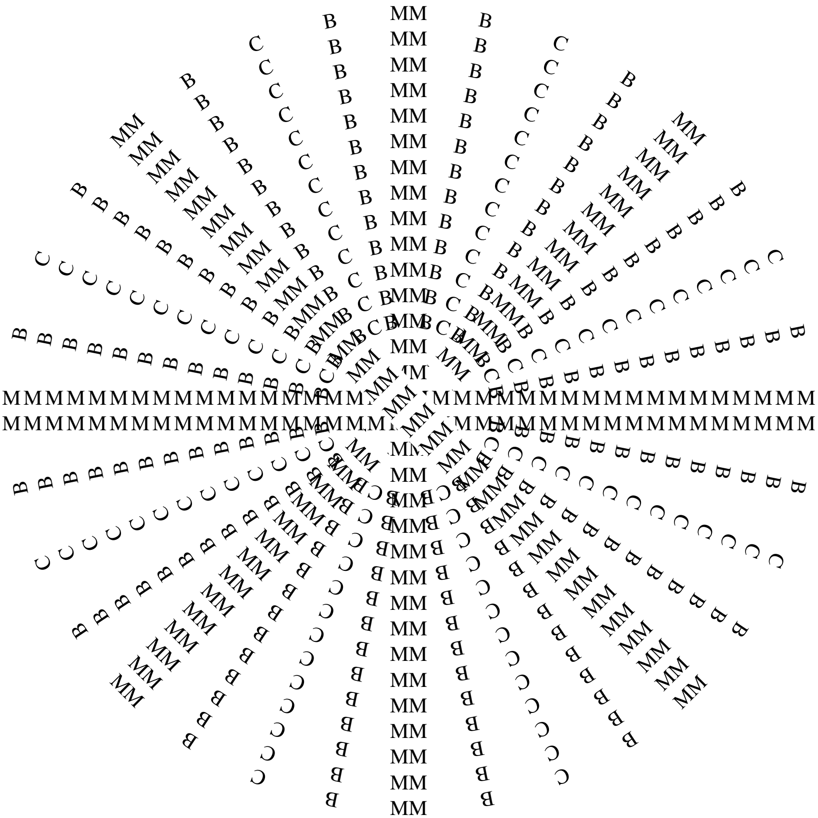

FIGURE 17.7. Okigbo circle design with a pair of M, maize, rows equally spaced and with increasing row spacing for C and B cultivars.

and the one in Figure 17.5 may both be included in the same e.u. by placing one plan in the top half and the other in the bottom half. This would allow for obtaining estimates of density and spacing effects in the same e.u. Alternatively, each plan could form an e.u. for one of the treatments in an experiment. The experimenter may change the above plans in any of a number of ways in order to study effects of interest.

One such alteration of the above circle designs is given in Figure 17.7. Here M, maize, has double rows equally spaced and with a pair of rows in a north-south direction and one pair in an east-west direction. Also, pairs of M rows are at a 45° angle to the previous pairs. The other cultivar rows are interspersed between these pairs of M rows. For this plan, the plants within a row are equally spaced, but this may be modified as described above. The density per hectare for C and B decreases as the distance away from the origin increases. This is also true for B and M and for C, B, and M, but not for the M pair of rows.

In addition to the preceding parsimonious plans, the snail designs of Volume I may be extended to include mixtures of *n* cultivars. For the three cultivars C, B,

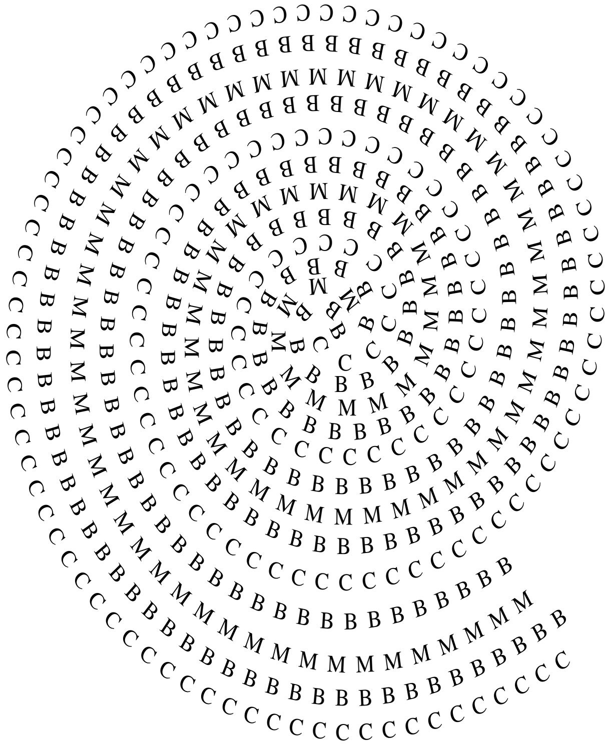

FIGURE 17.8. Snail design for three cultivars C, B, and M arranged in individual cultivar rows with equal spacing between plants within a row.

and M, a snail design with a row for each of the cultivars is presented in Figure 17.8. The plants within a row of a cultivar are equally spaced but may be altered to fit the objectives of the experimenter. The distance between the cultivar rows increases as the distance from the center increases. In this design, the widest spacing may be held to any desired distance. This is not possible for the fan or circle designs. A square-like area may be approached with the snail design, or parts of the design may be deleted to form a square e.u. These designs also allow orientation effects to be assessed.

For the plan in Figure 17.9, there is a single row arranged in a snail-like fashion. The cultivars appear in the row in the sequence of the arrangement. The plan used is a plant of C, a plant or hill of B, a plant or hill of M, and a plant or hill of B. The sequence is repeated throughout the row, winding in a snail-like fashion.

A snail design offers much more latitude in row spacings than does either the fan or circle designs. This advantage may be offset by the additional expertise in layout required over the fan or circle designs. For the latter, it is a simple matter to mark circles from the origin at all points where plants are to be planted. In Nigeria, Okigbo (1978) was able to use unskilled laborers to plant circle designs.

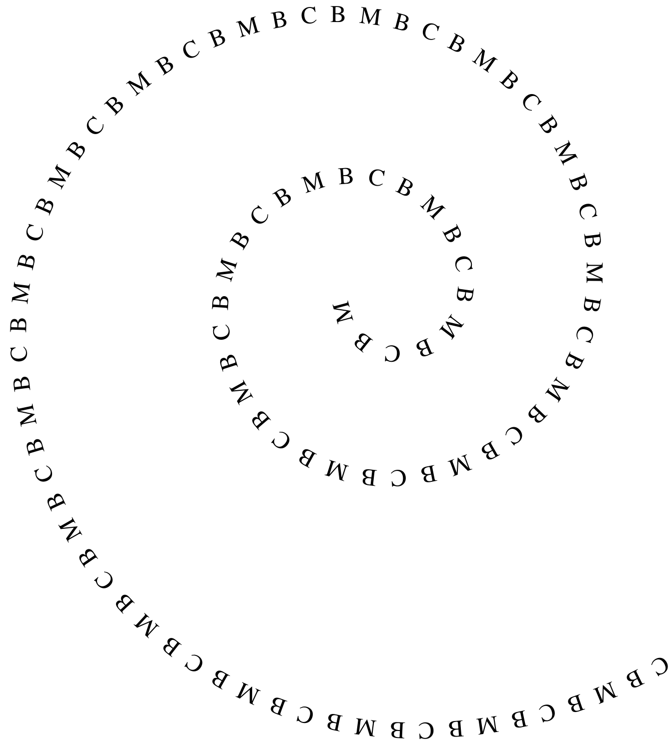

FIGURE 17.9. Snail design for three cultivars with the repeating sequence C, B, M, and B arranged in a single row. Distance between cultivars is a constant.

He assigned one person to each ray or row of the circle and all started planting their crop at the same time. Each person had instructions on the manner of planting the assigned crop. Each laborer planted only one crop in order not to burden them with too many instructions. It may be that a procedure can be devised to make the laying out of snail designs routine just as has been done for circle designs. The use of border material can be used to make the area for each e.u. a rectangular one.

17.4 Statistical Analyses for the Plans of Sections 17.2 and 17.3

For each density level and each spatial arrangement, the plant or group of plant yields are put on a yield per hectare basis. Then, these yields are plotted against density per hectare. The yield–density relation may have only a linear and a curvature component. A simple form of this is a quadratic regression equation of the form

$$Y_{ij} = \alpha + \beta_i d_{ij} + \gamma\, d_{ij}^2. \tag{17.1}$$

An estimate $\hat{d}_i$ of the maximum for this curve for cultivar i is $\hat{d}_i = -\hat{\beta}_i/2\hat{\gamma}_i$ where $\hat{\beta}_i$ and $\hat{\gamma}_i$ are least squares estimates of the parameters β_i and γ_i, respectively. These values are averaged over all e.u.s, yielding an estimate of the parameters. For the arrangements in Section 17.2, the density per hectare does not change. The yields per plant or group of plants need not be converted to yields per hectare, as they are all comparable and may be plotted directly against change in any of the factors. Analyses of variance may be performed on the intercepts, linear regression coefficients, quadratic regression coefficients, and/or the estimated maximum for the individual e.u.s. Analysis of orientation effects proceeds as described in Volume I. From the circle designs, the effects of north-south rows, east-west rows, northeast-southwest rows, and northwest-southeast rows can be ascertained. The effect of the various orientations on yield, insect level, disease level, and other factors can be studied.

Since the number of levels in the parsimonious plans can cover a wide range, the results will be useful in developing models for yield–density, yield–spatial, yield–intimacy, and yield–orientation relationships as well as multiple factor relations. Other factors such as rainfall, fertilizer, date of planting, cultivar, etc. may be used in conjunction with a parsimonious e.u. to develop yield–density and other yield relations in an efficient manner.

17.5 Discussion and Comments

The parsimonious designs discussed herein and the many variations possible offer an efficient tool for the experimenter to study the effects of spatial arrangement, density, intimacy, and orientation with a minimum expenditure of material, land area, and resources. They are especially useful in the preliminary investigation of the effects of these factors on responses for the cultivars in a mixture. Many levels of each of the factors may be included. After the results from investigations using the parsimonious plans, the experimenter may wish to switch to a factorial arrangement of levels around the estimated optimum for a more definitive study of the effects. The parsimonious plan may be viewed as a screening design to determine a neighborhood wherein the optimum level of each of the factors lies. Then, a more detailed investigation of this neighborhood may be investigated by again using parsimonious plans or a factorial arrangement.

During the preliminary stages of investigating effects of the factors, the experimenter may desire to use a single e.u. for a set of arrangements. These could then be included in an augmented experiment design (see Volume I) as the augmented treatments in an experiment with a set of standard replicated treatments. This would allow the screening of a large number of arrangements in an efficient manner. Each e.u. may require a border in order to eliminate border effects from the surrounding area. In Figures 17.3 and 17.4, e.g., a row of B, beans, could be used on the vertical and horizontal axes and the upper part could be filled out with

B to make a rectangular area for the e.u. The amount and nature of border material will depend on the experimental conditions present.

Figure 17.1 could be expanded to include spatial variation, density variation, and intimacy variation within the same e.u. or an e.u. for each could be included. From the preceding, it can be seen that there are many and diverse arrangements for studying spatial, density, intimacy, and orientation effects for mixtures of n cultivars. Row orientation in sugar beet production has been reported by Anda and Stephens (1996). Creativity on the part of experimenters will develop more and diverse arrangements in the spirit of parsimony. In screening genotypes as affected by various factors, use may be made of augmented designs and parsimonious arrangements to increase the efficiency of experimentation.

17.6 Problems

17.1 Construct arrangements of the form of those in Section 17.2 for $n = 4$ cultivars in the mixture.

17.2 Construct arrangements for $n = 4$ and 5 cultivars in a mixture similar to those in Figures 17.1 and 17.2.

17.3 Show how to adapt the Nelder (1962) designs in Figure 8.3, Volume I, for three and four treatments in a mixture.

17.4 Prepare figures comparable to Figures 17.3 and 17.4 for four cultivars in a mixture.

17.5 Prepare a figure comparable to Figure 17.5 for a mixture of four cultivars.

17.6 Prepare a figure comparable to Figure 17.5 with density per hectare constant and with four areas for three cultivars. (Note: Use the formula in Section 8.3, Volume I, to obtain areas of concentric circles such that areas between circles are equal.)

17.7 Construct a plan similar to Figure 17.1 in such a manner that row spacing, density, and intimacy effects are separately estimable within the same e.u.

17.8 Construct plans similar to Figure 5 of Mead and Riley (1981) for mixtures of three and four cultivars.

17.9 Detail a computational procedure for determining an optimal row spacing for your design in Problem 17.6.

17.10 Prepare a snail design for three cultivars in three separate rows and demonstrate how to estimate orientation effects.

17.7 Literature Cited

Anda, A. and W. Stephens (1996). Sugar beet production as influenced by row orientation. *Agron. J.* **88**, 991-996.

Anderson, D.A. and W.T. Federer (1975). Possible absolute determinant values for square (0,1) matrices useful in fractional replication. *Util. Math.* **7**, 135-150.

Federer, W.T. and K.E. Basford (1991). Competing effects designs and models for two-dimensional field arrangements. *Biometrics* **47**, 1461-1472.

Federer, W.T. and B. Scully (1992). Parsimonious statistical design and breeding procedure for evaluating and selecting desirable characteristics over environments. *Theoret. Appl. Genet.* **86**, 612-620.

Haizel, K.A. (1974). The agronomic significance of mixed cropping. *Ghana J. Agric. Sci.* **7**, 169-178.

Hiebsch, C.K., Fitetio-Kagho, A.M. Chirembo, and F.P. Gardner (1995). Plant density and soybean maturity in a soybean-maize intercrop. *Agron. J.* **87**, 965-969.

Kass, D.C.L. (1978). Polyculture cropping systems: Review and analysis. Cornell International Agriculture Bulletin 32, N.Y. State College of Agriculture and Life Sciences, Ithaca, N.Y.

Mead R. and J. Riley (1981). A review of statistical ideas relevant to intercropping research. *J. Roy. Statist. Soc., Series A* **144**, 462-509.

Nelder, J. (1962). New kinds of systemmatic designs for spacing experiments. *Biometrics* **18**, 283-307.

Okigbo, B. N. (1978). Personal communication.

CHAPTER 18

Some Analytical Variations for Intercropping Studies

18.1 Introduction

When systems, agricultural or otherwise, are the object of study in experiments, many and varied goals and analyses arise. In this chapter, some additional experimental variations and statistical analyses are described. The approaches in Volume I are expanded and extended to include more than two crops in a mixture. In the next section, a type of experimentation known as replacement series is discussed, wherein the proportions of the n crops in the mixture are varied in such a manner that the sum of the proportions is unity. The proportion is the population number for a sole crop required to plant that proportion of a hectare, say, occupied by a cultivar at the same rate as the sole crop. Graphical displays and numerical examples are included to illustrate various patterns of response and statistical analyses for replacement series experiments.

In Section 18.3, some combined yield comparisons are discussed. These are the mean method, half-hectare method, pure stand, production, land equivalent ratio, and total effective area. Numerical illustrations are included. In Section 18.4, five competition indices are described for n cultivars in the mixture. In the literature observed on these indices, it appears that the discussion has been for $n = 2$ only. Each of the indices are extended to include n cultivars. The indices are relative crowding coefficient, relative reproductive rate, competition index, coefficient of aggressivity, and competitive ratio. The linear programming techniques discussed in Volume I for two crops are extended to include n crops. A discussion of some of the aspects of this chapter is presented in Section 18.6. This is followed by a set of problems designed to illustrate and elucidate various aspects of this chapter.

18.2 Replacement Series

The larger the number n of cultivars in a mixture, the more complex is the construction of the *replacement series* treatment design as defined by Haizel (1974), in Volume I, and in the previous chapter. Likewise, the experiment becomes more complicated and difficult to conduct. The number of plants in a mixture for a cultivar is determined by its proportion in the mixture, p_i, and the population or number of plants used for the sole crop. The idea is to allot the proportions of crops in terms of the area they would occupy in the mixture; i.e., if the area for the mixture was subdivided so that the subdivisions were sole crops, the area for a crop would be its proportion in the mixture. For example, in a cassava–maize–bean mixture the sole crop densities might be 10,000, 40,000, and 120,000 plants per hectare, respectively. In a mixture with proportions 1:1:1, the mixture would contain 10,000/3 plants of cassava, 40,000/3 plants of maize, and 120,000/3 plants of bean. For a mixture of 1:2:2, there would be 2000 cassava plants, 8000 maize plants, and 120,000/5 = 24,000 bean plants in the mixture. A replacement series experiment is useful for studying the competitive effect relationships among the n cultivars in the mixture. It should be noted that the optimum proportion in a replacement series experiment may not be the population combinations producing maximum yields. The additive method described in the previous chapter may produce higher yields than any of the proportions in a replacement series.

The three patterns of response described by Willey (1979) and in Volume I for two cultivars may be used for polycultures of three or more cultivars. The *mutual cooperative* pattern of competition effects implies that all cultivars in the mixture perform better than expected on the basis of sole crop performance. In the *mutual inhibitive* pattern of competition effects, all cultivars perform poorer than would be expected on the basis of sole crop responses. In addition to mutually cooperative and inhibitive patterns, a third category is desirable, i.e., *neutral* competition effects wherein all cultivars in a mixture perform the same whether in a mixture or grown as a sole crop. Their responses are unaffected whether grown in polyculture or monoculture. For this case, the individual terms in a land equivalent ration (LER) would be $1/n$.

In the *compensation* pattern of responses, one or more of the cultivars have a cooperative effect and the rest have an inhibitive or neutral pattern of response. Considering the three patterns of responses described above, let + stand for cooperative pattern, – stand for inhibitive pattern, and 0 indicate a neutral pattern. Then, for $n = 3$ cultivars in the mixture, the possible combinations of compensatory and inhibitory patterns are

$$\begin{array}{cccccc} + & + & + & + & - & - \\ + & - & + & 0 & - & 0 \\ - & - & 0 & 0 & 0 & 0, \end{array}$$

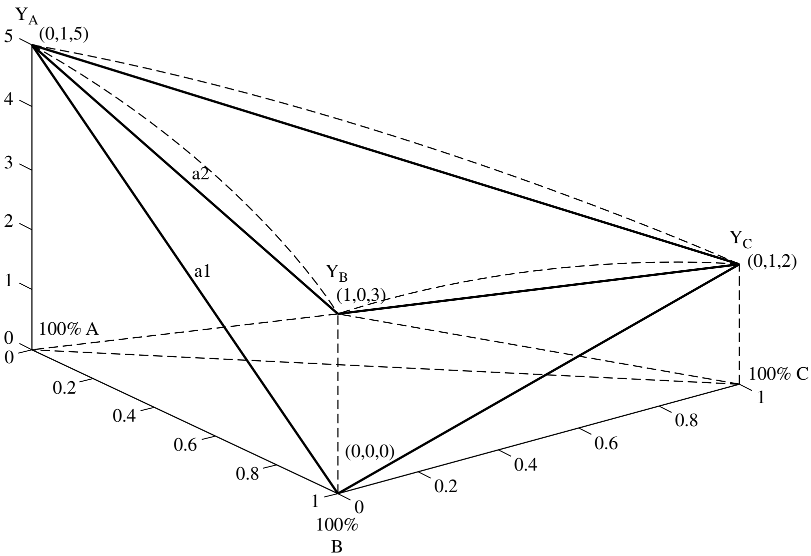

FIGURE 18.1. Mutual cooperation competition effects for three cultivars. Expected yield is on the triangle formed by the points Y_A, Y_B, and Y_C (sole crop responses). Observed yield is on the curved surface above the triangle $Y_A\ Y_B\ Y_C$.

for $n = 4$, the possible combinations are

+	+	+	+	+	+	−	−	−
+	+	−	+	+	0	−	−	0
+	−	−	+	0	0	−	0	0
−	−	−	0	0	0	0	0	0,

and for $n = 5$, the possible combinations are

+	+	+	+	+	+	+	+	−	−	−	−
+	+	+	−	+	+	+	0	−	−	−	0
+	+	−	−	+	+	0	0	−	−	0	0
+	−	−	−	+	0	0	0	−	0	0	0
−	−	−	−	0	0	0	0	0	0	0	0.

In general, there are $3(n-1)$ such patterns of compensatory and inhibitory effects.

A graphical representation of mutual cooperative competition effects is given in Figure 18.1. The equilateral triangle ABC formed by the points A, B, and C represents all possible combinations of mixtures of three such that the proportions add to one, i.e., $p_A + p_B + p_C = 1$. The triangle $Y_A\, Y_B\, Y_C$ formed by the points Y_A, Y_B, and Y_C represents the predicted responses based on sole crop yields, i.e., $p_A\, Y_A + p_B\, Y_B + p_C\, Y_C$. The curved surface above this triangle intersecting the dashed lines represents the observed responses for the mixture ABC. The

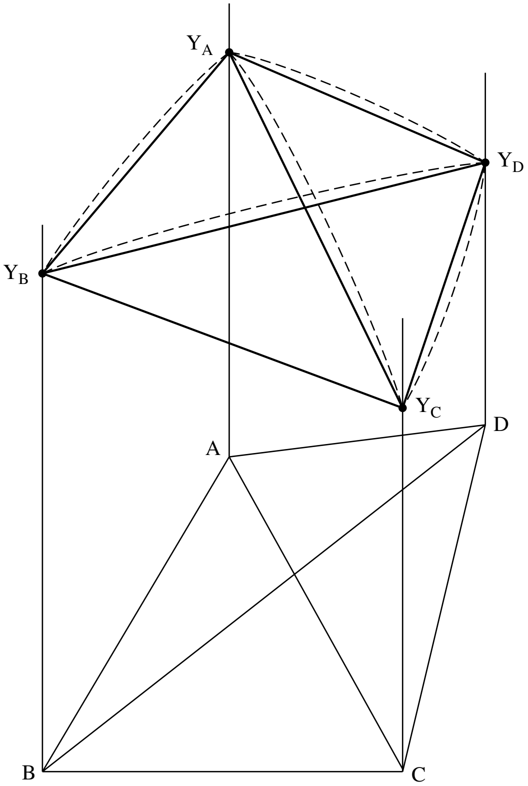

FIGURE 18.2. Mutual cooperation competition effects for four cultivars A, B, C, and D. Expected values are in the pyramid formed by the points Y_A, Y_B, Y_C, and Y_D. Observed responses are on the curved surface surrounding the pyramid Y_A Y_B Y_C Y_D.

quadrilateral Y_A Y_C AC formed by the points Y_A, Y_C, A, and C is for the two-crop mixture of A and C with $p_B = 0$ and with proportions of A and C varying from zero to 100%. Likewise, the quadrilaterals Y_A Y_B AB and Y_B Y_C BC are for mixtures of A and B, and B and C, respectively. The line $Y_A C$ represents the predicted sole crop yield for A with proportions from 100% of A, sole crop for A, to zero percent of A and 100% of C, sole crop for C. The dashed line above this line indicates the observed responses. Similar explanations are made for lines $Y_B A$, $Y_B C$, $Y_A B$, $Y_C A$, and $Y_C B$.

All possible combinations for $n = 4$ crops, A, B, C, and D, in a replacement series are obtained from $p_A + p_B + p_C + p_D = 1$, where $p_i, i =$ A, B, C, D, varies from zero to one. The totality of possible proportions is a pyramid with equilateral triangles on its four faces. This is depicted by the pyramid formed by the four points A, B, C, and D in Figure 18.2. The predicted or expected responses based on sole crop responses Y_A, Y_B, Y_C, and Y_D are given in the pyramid Y_A Y_B Y_C Y_D in the top part of Figure 18.2. When effects are mutually cooperative, the observed

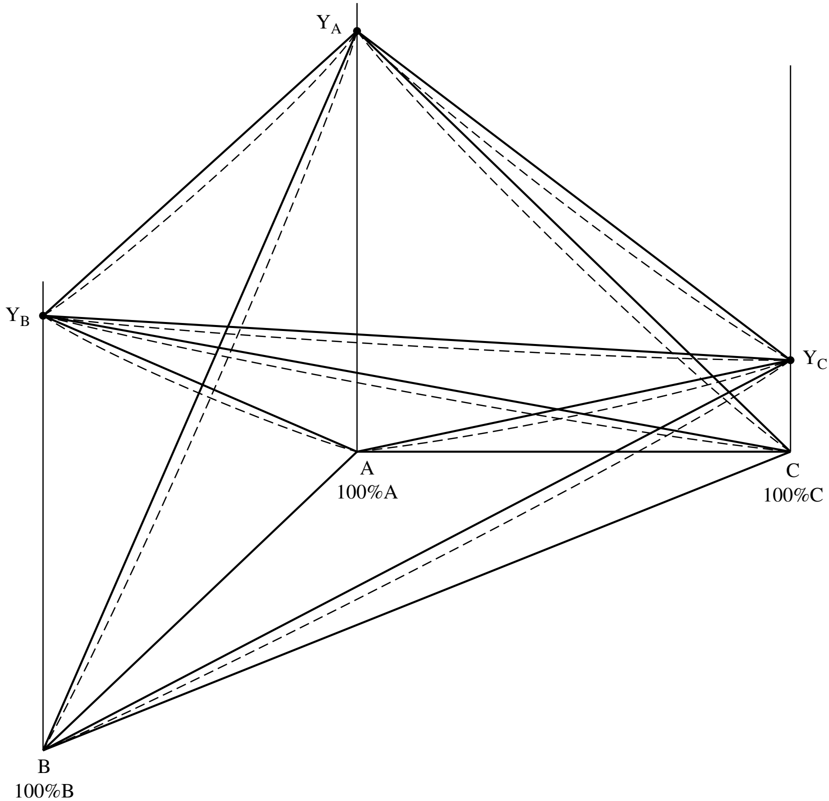

FIGURE 18.3. Mutual inhibition competition effects for three cultivars A, B, and C. Expected responses are on the triangle formed by the points Y_A, Y_B, and Y_C. Observed responses are on the curved surface below the triangle Y_A Y_B Y_C.

responses are indicated by a curved surface around the pyramid and intersecting the dashed lines. The additional details in Figure 18.1 may be included if desired.

To show mutually inhibitive competition effects, Figure 18.3 was prepared for $n = 3$ crops in a mixture. The graph is similar to Figure 18.1, except that the dashed lines are below rather than above their expected values. Neutral competition values would be exemplified by the predicted values from sole crop responses. For compensatory competition effects for $n = 3$ crops in the mixture, there are six cases to consider involving all three patterns of response. Graphs similar to Figures 18.1 and 18.3 may be used to depict the results for each of the six cases. Numerical examples showing the three types of response patterns are given in Examples 18.1 and 18.2.

There is a body of statistical literature on various kinds and properties of these mixture treatment designs such as used in replacemment series studies. A book on mixture designs by Cornell (1990) and one on optimal designs by Atkinson and Donev (1992) are suggested reading on these topics. A recent paper by Atkinson

TABLE 18.1. Yields = Y (g) and Predicted yields = P of All Possible Combinations of B, C, D, E, and F with Cultivar A. (Data Are from Table 13.3.)

Mixture		Culivar A	B	C	D	E	F	Total
ABC	Y	13.88	2.08	52.54				68.50
	P	22.58	1.64	39.54				63.76
ABD	Y	27.75	1.81		2.41			31.97
	P	22.58	1.64		2.71			26.93
ABE	Y	20.50	3.04			9.51		33.05
	P	22.58	1.64			8.96		33.18
ABF	Y	18.47	1.67				0.76	20.90
	P	22.58	1.64				1.59	25.81
ACD	Y	15.81		44.27	2.68			62.76
	P	22.58		39.54	2.71			64.83
ACE	Y	21.82		46.16		4.56		72.54
	P	22.58		39.54		8.96		71.08
ACF	Y	22.46		53.05			2.58	78.09
	P	22.58		39.54			1.59	63.71
ADE	Y	21.16			1.67	7.20		30.03
	P	22.58			2.71	8.96		34.25
ADF	Y	20.38			2.07		1.66	24.41
	P	22.58			2.71		1.59	26.88
AEF	Y	24.03				16.44	1.79	42.26
	P	22.58				8.96	1.59	33.13
Mean		20.66	2.15	49.01	2.21	9.43	1.70	

(1996) describes the usefulness of optimal designs. Reference to these citations should be made when planning replacement series type experiments and selecting the various p_i for the crops. Use of optimal designs makes for efficient experimentation. Assuming a quadratic relation for two crops A and B, a three-point optimal design would consist of the two sole crops and the 1:1 mixture. For three crops and a seven-point design, the ratios 1:0:0, 0:1:0, 0:0:1, 1:1:1, 1:1:0, 1:0:1, and 0:1:1 would suffice. The assumption of a quadratic surface may be erroneous, and the nature of the relationship will depend on the population densities used for the sole crops under investigation.

Example 18.1. The data from Table 13.3 may be analyzed in the mode of replacement series analyses. This is another example of the multiple goals and analyses that are associated with experiments on polyculture systems and sole crops. The cooperative, inhibitive, and neutral effects of the cultivars in the mixture may be examined. For this example, the 10 three-cultivar combinations involving cultivar A are used. The individual cultivar and the mixture totals of 3 replicates for the 10 mixtures are presented in Table 18.1. The results in Table 13.3 are for a four-plant total. Therefore, the yield of any cultivar in the mixture may be compared directly with its sole crop yield, which is also on a four-plant basis. The three-crop mixture total is on a 12-plant basis, 4 from each cultivar. The sum of the 3 corresponding sole crop yields in a mixture is also on a 12-plant basis.

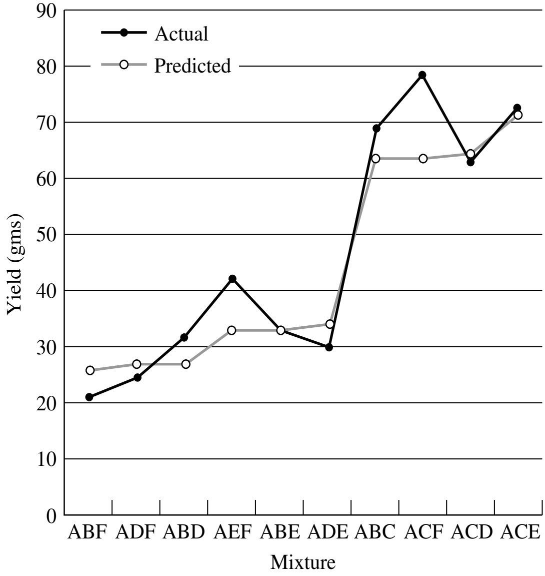

FIGURE 18.4. Yields (m) and predicted yields (s) for mixtures arranged in increasing order of predicted yields for a mixture. Data from Table 18.1.

Cultivar A exhibited a cooperative effect only for mixtures ABD and AEF. In mixtures ABC, ABF, and ACD, A showed an inhibitive effect. The effect was neutral or slightly inhibitory in the remaining five mixtures where A appeared. Cultivar B showed a cooperative or neutral effect in the four mixtures where it appeared. Cultivar C showed rather large cooperative effects in all four mixtures. Cultivar D showed neutral to inhibitory effects. Cultivar E had two cooperative and two inhibitive responses in the four mixtures. Cultivar F showed three cooperative responses and one inhibitive response in the four mixtures containing F.

A graphical representation of results is given in Figure 18.4. The mixture designations on the abscissa are equally spaced and are arranged in increasing order of sole crop predicted yields. If the mixture designations on the abscissa had been spaced according to the mixture totals, the predicted totals would have been on a line with a slope of 45°. Graphical presentations such as Figure 18.4 may be more informative than the last column of Table 18.1.

So far, nothing has been said about the statistical significance of the cooperative or inhibitive competition effects of the cultivars in the 10 mixtures. Using a simple nonparametric procedure like chi square, it may be noted that in the 10 cases for cultivar A, 8 were less than the sole crop, giving a chi square value of 3.6 versus the tabulated 5% value of 3.84. For cultivars B and C, all four yields are above the sole crop value, giving a chi square value of four. From Table 13.5, the various standard errors of a difference are computed in the following manner:

$$\text{Cultivar A}: \quad \sqrt{2(3)(7.738)} = 6.81 \text{ for mixture total versus sole crop}$$

$$\sqrt{7.738(3/10+3)} = 5.05 \text{ for mean of 10 mixtures versus sole crop}$$

$$\text{Cultivar B}: \quad \sqrt{2(3)(0.105)} = 0.79 \text{ for mixture total versus sole crop}$$

$$\sqrt{0.105(3/4+3)} = 0.63 \text{ for mean of four mixtures versus sole crop}$$

$$\text{Cultivar C}: \quad \sqrt{2(3)(34.62)} = 14.41 \text{ for mixture total versus sole crop}$$

$$\sqrt{34.62(3/4+3)} = 11.39 \text{ for mean of four versus sole crop}$$

$$\text{Cultivar D}: \quad \sqrt{2(3)(0.157)} = 0.97 \text{ for mixture total versus sole crop}$$

$$\sqrt{0.157(3/4+3)} = 0.77 \text{ for mean of four versus sole crop}$$

$$\text{Cultivar E}: \quad \sqrt{2(3)(3.159)} = 4.35 \text{ for mixture total versus sole crop}$$

$$\sqrt{3.159(3/4+3)} = 3.44 \text{ for mean of four versus sole crop}$$

$$\text{Cultivar F}: \quad \sqrt{2(3)(0.096)} = 0.76 \text{ for mixture total versus sole crop}$$

$$\sqrt{0.096(3/4+3)} = 0.60 \text{ for mean of four versus sole crop.}$$

Since the mixture values in Table 18.1 are totals of three replicates, this is the 3 that appears in the standard error of differences between a mixture value and a sole crop value. The 2 is because this is a difference being compared. Each cultivar has its own variance. The difference between cultivar A total in mixture ABC and the sole crop is $13.88 - 22.58 = -8.70$ with a standard error of difference of 6.81. This gives a t value of 1.28 and a probability of a larger t of about 0.20. The difference between cultivar C total in mixture ABC and sole crop C is $52.54 - 39.54 = 13.00$ with a standard error of a difference of 14.41. The difference between cultivar B total in mixture ABE and sole crop B is $3.08 - 1.64 = 1.44$ with a standard error of a difference of 0.79. The corresponding t value is 1.82 with a probability of a larger t value of about 0.07. The difference between Y and P for mixture ACF total and sole crop totals is $78.09 - 63.71 = 14.38$ with a standard error of a difference of $\sqrt{3(2)(7.738 + 34.62 + 0.096)} = 15.96$.

Example 18.2. The data from Example 13.2 are used to exemplify the analysis of a replacement series experiment. A genotype of castrol bean, one of maize, and one of bean were used to form the mixtures and proportions of two and of three cultivars. The sole crops were also included. The observed yields of individual cultivars and of mixture totals are presented in Table 18.2 along with sole crop yields and sole crop predicted replacement series yields for each mixture. A graphical presentation

TABLE 18.2. Yields (kg/ha) of Castrol Bean (C), Maize (M), and Bean (B) and Replacement Series Predicted (Pred.) Yields Based on Sole Crop Yields. (Data from Example 13.2. Means are for Mixtures.)

	Castrol bean = C		Maize = M		Bean = B		Total	
Mixture	Yields	Pred.	Yield	Pred.	Yield	Pred.	Yield	Pred.
1C:1M	1080	936	2170	1084			3250	2020
1C:2M	728	624	2290	1445			3018	2069
1C:3M	847	468	1936	1625			2783	2093
1C:2B	1148	624			1005	671	2153	1295
1C:3B	888	468			1150	755	2038	1223
1C:4B	746	374			858	806	1604	1180
1C:2M:1B	561	468	1991	1084	92	252	2644	1804
1C:1M:2B	1242	468	1449	542	540	504	3231	1514
1C:2M:2B	697	374	1657	867	340	403	2694	1644
C (sole)	1871	1871					1871	1871
M (sole)			2167	2167			2167	2167
B (sole)					1007	1007	1007	1007
Mean	882	534	1916	1108	664	565	2602	1649

of the results is given in Figure 18.5. The predicted yield for the mixture 1C:2M:1B, for example, is 1871/4 + 2(2167)/4 + 1007/4 = 468 + 1084 + 252 = 1804, and the others are similarly computed.

In every case except two, the yields exceed those predicted upon the basis of sole crop yields. For these particular cultivars, population densities, intimacy relations, and spatial arrangements, polyculture was very beneficial. Maize did especially well in that monoculture yields were equaled or almost equaled in the four mixtures 1C:1M, 1C:2M, 1C:3M, and 1C:2M:1B. Except for the two mixtures 1C:2M:1B and 1C:2M:2B, beans did better than expected for mixtures where it occurred. Castrol beans had approximately double the predicted replacement series yields for the mixtures 1C:3M, 1C:2B, 1C:3B, 1C:4B, 1C:1M:2B, and 1C:2M:2B. The castrol bean yields were higher than expected for 1C:1M, 1C:2M, and 1C:2M:1B, but the yields were not double the predicted, as in the other mixtures.

A graphical presentation of the data in Table 18.2 may be made as in Figure 18.5. The equilateral triangle MBC formed by the points M, B, and C represents all possible replacement series proportions for these cultivars. The triangle formed by the points Y_M, Y_B, and Y_C represent the predicted yields for these proportions. The lines through the points l, m, and n are on the plane forming the triangle $Y_M Y_B Y_C$, and they are the expected replacement series yields based on the sole crop yields. The line ho including the points h and o is the observed response for the mixture 1C:2M:2B. The line ik is the observed response for the mixture 1C:2M:1B. The line qj is the observed response for the mixture 1C:1M:2B. Note that the line ik ends before it reached the plane $Y_M Y_B Y_C$; i.e., the observed response is less than the expected calculated from sole crop yields.

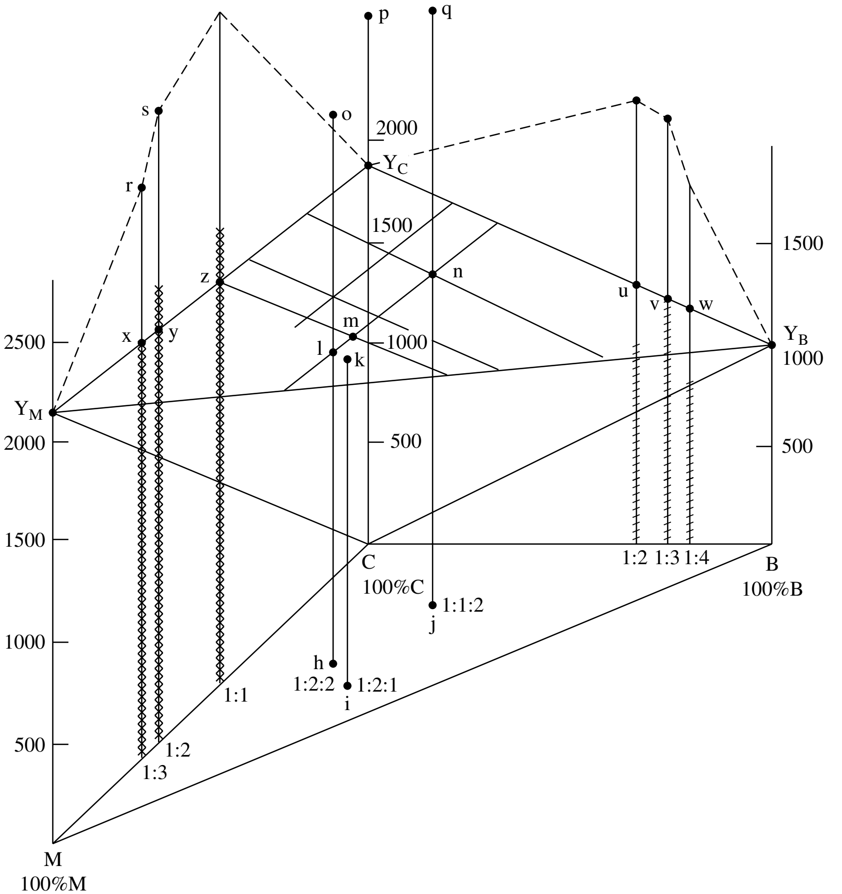

FIGURE 18.5. Graphical representation of cooperative and inhibitive effects in a replacement series for the data of Example 18.2.

Mixtures of two cultivars may be included in the three-dimensional graph in Figure 18.5. In the quadrilateral formed by the points Y_M, Y_C, M, and C, the predicted values of maize and castrol bean combinations in all proportions are shown. The line $Y_M Y_C$ is the predicted replacement series yields. The line rx containing the points r and x is the observed response for the mixture 1C:3M. The line sy represents the observed yield for the mixture 1C:2M, and the line tz displays the observed yield for the mixture 1C:1M. The bottom part of the lines with x's on the spike designates the observed yield of maize and the top part represents the yield of castrol beans in the mixture. The line $Y_M C$ represents the predicted yields for all proportions of maize ranging from 100% to zero. Note that the observed maize yields are considerably above this line, indicating the cooperative competition effect for maize.

For the two-crop mixture of castrol beans and beans, the quadrilateral $Y_C Y_B BC$ formed by the points Y_C, Y_B, B, and C is used to present the results. The line Y_B to C represents the expected bean values from 100% bean to 0% bean in the two-crop mixture. The observed bean yields are denoted by putting dashes on the spikes going through the points u, v, and w. Note that all exceed the predicted yields on the line $Y_B C$, indicating the cooperative competition effect of bean in the two-crop mixtures with castrol beans.

Since the mixture yield of castrol beans is placed on the top parts of the spikes representing observed yields in Figure 18.5, the predicted yields for castrol beans in two-crop mixtures were not included. These would be indicated by the two lines $Y_C M$ and $Y_C B$. They could be included in the graph with additional spikes in addition to the corresponding maize and bean spikes already shown in Figure 18.5.

The dashed lines at the top of the graph represent the observed yields from all possible proportions of maize and castrol bean, and bean and castrol bean mixtures. The true picture is more likely a smooth surface rather than the line-segment depiction. Similarly, dashed and curved lines could be drawn from the observed three-crop mixture yields to the sole crop and two-crop mixture observed yields. The resulting curved three-dimensional surface would depict yields for all possible combinations of two and three crops that would result from the patterns of responses found for these three crops.

Individual error variances for each of the three crops were not available. When such error variances are available, the procedure given in Example 18.1 may be followed to assess statistical significance. Using a nonparametric procedure like chi square, highly significant chi-square values were obtained for castrol beans, nine out of nine cases in which predicted yields were exceeded, and for maize, six out of six cases in which predicted yields were exceeded. For bean mixtures, bean yield exceeded predicted yield in four out of the six cases, resulting in a chi-square value of 0.67. All of the nine mixture totals exceeded the predicted totals, resulting in a chi-square value of 9.00. Which mixture or mixtures would be selected for use depends on the goals of the user. Mixture 1C:1M:2B produced more than double expected on the basis of sole crop yields. Over all nine mixtures for castrol beans, the ratio of the mixture mean to the mean of the predicted values is 882/534 = 1.65, which is 65% more than when grown as a sole crop. Maize produced 73% more in polyculture than in monoculture, and beans produced 18% more. The mixture mean for the polyculture was 51% larger than the monoculture. Thus, the cooperative competition effects of castrol beans and maize are sizable. The three-crop mixtures which included 2M had a depressing effect on bean yields, especially in the mixture 1C:2M:1B.

18.3 Combined Yield Comparisons

18.3.1 Mean Method

Perhaps the simplest method for comparing monoculture with polyculture responses is the method of comparing the mean yield of the mixture with the mean of the mean yields of the cultivars in the mixture, or the *mean method* . This procedure was suggested by Donald (1963) as a method for evaluating polyculture responses for hay crops and mixed cereal cultivars. To illustrate, from Example 13.3, the yield of the maize and bean monocultures at Caruara were 2451 and 69 kg/ha, respectively. The mean of the monocultures is 1260. The yields from the maize–bean mixture were 2599 and 30, respectively, with mixture mean of 2599 + 30 = 2629. As Kass (1978) and others point out, a mean yield of cultivar yields from cultivars as different as these is a meaningless concept, although it could be appropriate for hay, mixtures of cereals, and mixtures of cultivars of the same species.

The method extends immediately to n cultivars in a mixture and for different proportions of cultivars in the mixture. From Example 13.2, the polyculture yields of castrol beans, maize, and beans in the ratio 1:2:1 were 561, 1991, and 92, respectively. The monoculture yields of castrol beans, maize, and beans were 1871, 2167, and 1007, respectively. The weighted mean of the monocultures yields is (1871 + 2(2167) + 1007)/4 = 1803. The mixture yield is 561 + 1991 + 92 = 2644.

18.3.2 Half-Hectare Method

The inadequacy of the preceding method was recognized by Pilz (1911) and Lipman (1912) much earlier. Their suggestion was to use the yield per unit area of the two crops in the mixture. Their *half-hectare* method for the data of Example 13.3 above is 2451/2 = 1225.5 for maize and 69/2 = 34.5 for beans. For maize, the comparison is 1225.5 for sole versus 2599 for the mixture. For beans, the comparison is 34.5 for sole versus 30 for the mixture of maize and beans. Morrish (1934) and Willey and Osiru (1972) criticized the method for assuming only equal proportions of areas for the two crops in the mixture.

For mixtures of size n, the half-hectare name could be changed to $1/n$**th hectare**, as it may be easily extended. Also, there appears to be no reason for assuming equal proportions of the crops in the mixture except for the name half-hectare. For the castrol bean, maize, and bean mixture from Example 13.2 in the ratio 1:2:1, maize occupies one-half of the area and the other two crops one-fourth of the area. The comparison for castrol bean sole crop one-fourth mean is 1871/4 = 467.75 versus 561 for the mixture yield. For maize, the half-hectare sole yield is 2167/2 = 1083.5 versus 1991 for the mixture. For beans, the one-fourth hectare sole yield is 1007/4 = 251.75 versus 92 for the bean yield in the mixture. Maize and castrol yields were increased using this mixture, but bean yield was reduced.

18.3.3 Pure Stand Production

Morrish (1934) suggested taking the yield of crop i in a mixture, Y_{mi}, and dividing it by the pure stand yield, Y_{si}, i.e., $Y_{mi}/Y_{si} = L_i$, and then summing the two fractions obtained to obtain what is now known as a land equivalent ratio, LER, or a relative yield total, RYT. We denote this method as a *pure stand production*, PSP. The yield of the mixture is $Y_{m1} + Y_{m2}$ and dividing the mixture yield by the sum of the fractions, i.e., $(Y_{m1} + Y_{m2})/(L_1 + L_2) = (Y_{m1} + Y_{m2})/\text{LER}$. This gives the percentage benefit of the polyculture over the monoculture. For the maize–bean data from Example 13.3, the yield of the mixture is 2629 and the LER is 2599/2451 + 30/69 = 1.06 + 0.43 = 1.49. Thus, the Morrish measure of pure stand estimated yield is 2629/1.49 = 1764, which is the yield from one hectare grown in the relative proportions of 1.06/1.49 = 0.71 of maize and 0.43/1.49 = 0.29 of beans. The efficiency of polyculture to monoculture for this mixture is 2629/1764 = 1.49, or an increase in efficiency of 49%. Note that this is the same efficiency obtained from the LER. Hence, the step of dividing the mixture total by an LER is unnecessary.

To obtain a PSP, Willey and Osiru (1972) suggest taking L_1/LER hectare as the area of sole crop one required to produce as much as was obtained from the mixture and L_2/LER hectare as the area of sole crop two required to produce as much as was obtained from the mixture. For the above example, $L_1 = 1.06$ and $L_2 = 0.43$. A hectare divided in the proportions indicated from the mixture yields would result in $2451(1.06/1.49) = 1744$ kg for maize and $69(0.43/1.49) = 20$ kg for beans, or a total of 1764 kg/ha.

Extending the PSP to more than two crops is straightforward. Consider the data from Example 13.2 described above. A LER $= L_c + L_m + L_b = 561/1871 + 1991/2167 + 92/1007 = 0.300 + 0.919 + 0.091 = 1.310$, $L_c/\text{LER} = 0.300/1.310 = 0.229$, $L_m/\text{LER} = 0.919/1.310 = 0.702$, and $L_b/\text{LER} = 0.091/1.310 = 0.069$. Thus, 70.2% of the hectares would need to be planted to maize, 22.9% to castrol beans, and 6.9% to beans in order to obtain the same proportion of yields of the three crops as obtained from the mixture. The $\text{PSP} = 0.229(1871) + 0.702(2167) + 0.069(1007) = 2019$. The efficiency of this mixture relative to sole cropping is 2644/2019 = 1.31, which is the LER.

18.3.4 Relative Yield Total or Land Equivalent Ratio

A relative yield total, RYT, was defined by de Wit and van den Bergh (1965) and van den Bergh (1968) primarily in the context of replacement series and for the proportions of crop yields in mixtures to sole crop yields. A land equivalent ratio, LER, which is the same as RYT, was formally defined by Willey and Osiru (1972) according to Mead and Riley (1981) and by IRRI (1974) according to Kass (1978). The Spanish equivalent of LER is *Uso Equivalente de Tierra*, UET (see Soria *et al.*, 1975). Morrish (1934) and Niqueux (1959) had used this index much earlier. An LER is now commonplace in literature on the analysis of intercropping experiments. This measure is the one discussed in Chapter 13.

18.3.5 Total Effective Area

This is a measure to determine the utility of mixed cropping and sole cropping to achieve a desired goal. This measure was described in Volume I for two crops. The measure may be extended to three crops, say, as follows. Let A_i, $i = 1, 2, 3$, be the area for sole crops one, two, and three, respectively, and let A_m be the area devoted to the three-crop mixture, then the total effective area (TEA) is

$$\begin{aligned}\mathrm{TEA} &= A_1 + A_2 + A_3 + \mathrm{LERA}_m \\ &= A_1 + A_2 + A_3 + (L_1 + L_2 + L_3)A_M. \qquad (18.1)\end{aligned}$$

The L_i are defined above. The extension to n crops is straightforward as long as only the sole crop mixtures and the n cultivar mixture are used. If sole crops, two-crop mixtures, and three-crop mixtures are to be used as was done in Example 18.2, then

$$\begin{aligned}\mathrm{TEA} = {} & A_1 + A_2 + A_3 + \mathrm{LER}_{12}A_{12} + \mathrm{LER}_{13}A_{13} \\ & + \mathrm{LER}_{23}A_{23} + \mathrm{LERA}_m, \qquad (18.2)\end{aligned}$$

where LER_{12} is a land equivalent ratio for the two-crop mixture of cultivars 1 and 2, and A_{12} is the area devoted to growing this two-crop mixture. The other terms are defined similarly. The extension to n crops with all sizes of mixtures from 2 to n is straightforward even if tedious.

The above index may be altered to take into account the proportions of crops desired. Let the proportions be $\lambda_1{:}\lambda_2{:}\lambda_3$ such that $\lambda_1 + \lambda_2 + \lambda_3 = 1$ and let q_i be the amount of seed sown per unit area for cultivar i, $i = 1, 2, 3$. The quantity of seed sown for the three-crop mixture is $q_m = \lambda_1 q_1 + \lambda_2 q_2 + \lambda_3 q_3$. Then, for equation (18.1),

$$\begin{aligned}\mathrm{TEA} = {} & A_1 + A_2 + A_3 + (q_1 L_1/\lambda_1 + q_2 L_2/\lambda_2) \\ & + (q_3 L_3/\lambda_3)\, A_m/3q_m. \qquad (18.3)\end{aligned}$$

TEAs for additional sized mixtures in addition to the n mixture may be developed as described above.

18.4 Competition Indices

18.4.1 Relative Crowding Coefficient

The relative crowding coefficient, RCC, was presented by de Wit (1960) and further developed by Hall (1974a,b). A coefficient K_1 is defined for crop one in a mixture of crops one and two as

$$K_1 = Y_{m1}p_2/p_1(Y_{s1} - Y_{m1}) = L_1 p_2/p_1(1 - L_1), \qquad (18.4)$$

where p_1 is the sown proportion of cultivar 1, $p_2 = 1 - p_1$, and the other symbols are as defined above. Instead of the sown proportion, the proportion of numbers of plants emerged or even number of plants at harvest could be used as well. For competition studies, the number of plants at a given stage would appear to be more appropriate data than number of seeds sown, since different cultivars may have different emergence rates. A value of $K_i > 1$ indicates that cultivar i has done better than expected based on sole cultivar results, and a value of $K_i < 1$ indicates that the cultivar has done poorer than expected. For a mixture of two cultivars, RCC is defined to be

$$\mathrm{RCC} = K_1 K_2 = L_1 L_2/(1 - L_1)(1 - L_2). \tag{18.5}$$

For n greater than two cultivars in a mixture, K_i is defined to be

$$K_i = Y_{mi}(1 - p_i)/p_i(Y_{si} - Y_{mi}) = L_i(1 - p_i)/p_i(1 - L_i). \tag{18.6}$$

RCC is defined to be

$$\mathrm{RCC} = \Pi_{i=1}^{n} K_i = \Pi_{i=1}^{n} L_i(1 - p_i)/p_i(1 - L_i). \tag{18.7}$$

The quantity $1 - p_i$ is the space not occupied by cultivar i, which is p_2 when n equals 2. As stated above, the RCC could be calculated for one or more of the following stages:

(i) number of seeds sown per cultivar,
(ii) number of plants emerged per cultivar, and/or
(iii) number of plants producing responses Y_{mi} and Y_{si} at intervals during the growing season.

The last one may be the appropriate one for agronomic investigations. In place of actual number of plants, the sowing rate used for each sole crop may need to to be considered. For example, maize may be planted at the rate of 40,000 plants per hectare and beans at 120,000 plants per hectare. The proportion of the area occupied by a cultivar must be considered in studies of this nature.

18.4.2 Relative Reproductive Rate

De Wit (1960) defined another quantity called *relative reproductive rate*, RRR, as the ratio of the proportion of cultivar 1 in the harvested mixture to the proportion of cultivar 1 in the sown (emergence) mixture to the ratio of the proportion of cultivar 2 in the harvested mixture to the proportion of cultivar 2 in the sown (emergence) mixture. Symbolically, for crop one this is

$$\mathrm{RRR}(1) = \frac{(Y_{m1}/N_1)}{(Y_{m2}/N_2)} = \mathrm{RCC}(1)\left(\frac{Y_{s1}}{Y_{s2}}\right), \tag{18.8}$$

where N_i is the number of plants of cultivar i. Note that N_i may be replaced by $N_i/(N_1 + N_2)$ or by the proportion of area occupied by cultivar i in the formula for RRR.

We extend RRR(1) in equation (18.8) to cover the case of cultivar i in a mixture of n cultivars as follows:

$$\begin{aligned}
\text{RRR(i)} &= (Y_{mi}/N_i)\left(\sum_{j=1,\neq i}^{n} Y_{mj}/N_j\right)^{-1}(n-1) \\
&= \qquad K_i Y_{si} \Big/ \sum_{j=1,\neq i}^{n} \frac{Y_{sj}}{(n-1)}. \qquad (18.9)
\end{aligned}$$

The yield of a cultivar is proportional to the relative space it occupies. De Wit (1960) stated that in certain cases with oat and barley mixtures, polyculture would be beneficial for the following:

(i) The 1000 kernel weight of oats was greater when intercropped with barley because the earlier maturing barley provides less competition to oats in the later developmental stages.
(ii) Oats reduced lodging in barley.
(iii) In a field heterogeneous for pH, barley and oats would both be located in low- and high-pH areas, where each was better adapted.
(iv) Elimination of one cultivar by disease or insects would allow a proportionately larger area than expected on the basis of the number of seeds sown or plants emerged.

In the last situation, de Wit (1960) stated the RCC of the resistant cultivar would be increased. It is stated in Kass (1978) that de Wit's model does not hold when the space shared by the two cultivars in a mixture is not the same as that available for monocultures—a reason frequently cited for the increased production often obtained with polycultures. He further says that polyculture is advantageous if RCC is greater than unity. De Wit (1961) extended his model for three cases where it did not hold as follows:

(i) The RRR is 1 and the proportion of cultivars harvested is the same as that seeded. In this case, the log of the ratio of cultivars in the harvested mixture is plotted against that of the cultivars in the sown mixture, and a straight line with unit slope and passing through the origin results.
(ii) RCC is not equal to unity because one cultivar has a competitive advantage over the other when competing for the same space. Here, the plot of the ratios of harvested yields against the seeded mixtures is a straight line not passing through the origin. This situation occurs when the two cultivars growing simultaneously have different growth curves.
(iii) RCC is not equal to 1 because the cultivars together exploit a larger or smaller share of the environment than either cultivar grown as a sole crop. If one cultivar benefits from the presence of the other so that the two cultivars are not competing for the same space, the plots of the ratios of harvested and sown mixtures will have a slope less than 1 and will not pass through the origin. If one cultivar has a harmful effect on another, the slope of the ratios

will have a slope greater than unity and will not pass through the origin. In both cases, the curves tend to be S-shaped rather than straight lines.

18.4.3 Competition Index

Donald (1963) introduced an equivalence factor for cultivars in terms of the number of plants required to produce a stated yield per unit area using N_{si} sole crop i plants and N_{mi} mixture plants. His *competition index* [developed by C. A. MacIntyre according to Kass (1978)] is defined as

$$\mathrm{CI} = (N_{s1} - N_{m1})(N_{s2} - N_{m2})/N_{m1}N_{m2}. \tag{18.10}$$

N_{si} is determined from a yield–density relationship as the number of sole crop i plants needed to produce the specified yield obtained from the mixture with N_{mi} plants. In order to obtain a yield–density relationship, several densities of sole crop i are required. Kass (1978) and Mead and Riley (1981) and others have criticized the method for requiring too many sole crop experimental units. It should be noted that N_{si} may be estimated as follows from one experimental unit. The yield per plant for a sole crop is Y_{si}/N_{si} for crop i and Y_{mi}/N_{mi} for mixture crop i. Then, N_{si} is estimated as

$$N_{si}^* = N_{mi}Y_{si}/Y_{mi}. \tag{18.11}$$

When $N_{si} = 2N_{mi}$, the competition index (CI) equals 1. When CI is less than unity, a beneficial association is indicated. If CI is greater than 1, a harmful association is indicated. For $Y_{si}/N_{si} = Y_{mi}/N_{mi}$ and $N_{si} = 2N_{mi}$, the LER $= 1.0 + 1.0 = 2$. If any of the $N_{si} = N_{mi}$, CI is equal to zero.

In light of the above, a competition index could be defined as

$$\mathrm{CI}^* = (Y_{s1}/N_{s1} - Y_{m1}/N_{m1})(Y_{s2}/N_{s2} - Y_{m2}/N_{m2}). \tag{18.12}$$

This form of the index compares the per plant yields of sole with mixture for cultivar i. If the per plant yields are equal for either (or both) cultivar(s), the index is zero. A positive value of the index for crop i would indicate a harmful effect and a negative value a beneficial effect for the crop i mixture. CI* [eq. (18.12)] would be negative when one of the two terms is negative and the other positive, indicating a differential effect for the two cultivars in the mixture. Also, CI* would be zero if one or the other of the terms is zero. Thus, valuable information is contained in the individual terms of the competition index.

In discussing CI, Kass (1978) states that it was designed for use in pasture species studies where the yield–density relationship is less complex than for other cultivars but was used by several authors for mixtures of annual crops. These authors did not appear to have constructed a yield–density relation for a crop over a range of sole crop densities. Furthermore, they obtained CI values less than 1 where the polyculture did not appear beneficial by other determinations, thus casting doubt on their use and application of the method to annual crops. When using CI, it is important to have comparable populations in sole and in mixed stands.

A competition index for a mixture of n cultivars is

$$\mathrm{CI} = \sum_{i=1}^{n} (N_{si} - N_{mi})/N_{mi} \tag{18.13}$$

or

$$\mathrm{CI}^* = \Pi_{i=1}^{n} (Y_{si}/N_{si} - Y_{mi}/N_{mi}). \tag{18.14}$$

Here, again, the individual terms of the index need to be examined to ascertain the competitive nature of each cultivar in the mixture.

18.4.4 Coefficient of Agressivity

To study the dominance of one cultivar over another, a *coefficient of aggressivity*, COA, was proposed by McGilchrist and Trenbath (1971) for replacement series studies. The coefficient is defined in terms the relative yield increase of cultivar 1 to cultivar 2 and is

$$\mathrm{COA(1)} = Y_{m1}/Y_{s1}p_1 - Y_{m2}/Y_{s2}p_2 = L_1/p_1 - L_2/p_2. \tag{18.15}$$

Mead and Riley (1981) state that there is difficulty in interpreting COA in that apparent dominance is related to the particular densities used for the sole crop yields. Dominance patterns can be reversed by changing sole crop densities.

For n cultivars in a mixture, the dominance is defined in terms of the dominance of cultivar i to the mean of the remaining cultivars in the mixture

$$\mathrm{COA(i)} = L_i/p_i - \left(\sum_{j=1,\neq i}^{n} L_j/p_j \right) /(n-1). \tag{18.16}$$

18.4.5 Competitive Ratio

As an alternative to COA, Willey and Rao (1980) proposed the following *competitive ratio*, CR, for the dominance of cultivar 1 relative to cultivar 2

$$\mathrm{CR(1)} = \frac{(Y_{m1}/Y_{s1}p_1)}{(Y_{m2}/Y_{s2}p_2)} = (L_1/L_2)(p_2/p_1). \tag{18.17}$$

For n cultivars in the mixture, the dominance is defined in terms of the dominance of cultivar i to the mean of the remaining $n-1$ cultivars in the mixture

$$\mathrm{CR(i)} = (L_i/p_i) \left(\sum_{j=1,\neq i}^{n} L_j/p_j \right)^{-1} (n-1). \tag{18.18}$$

Willey (1979) compared the RCC, LER, and COA using data from an experiment involving all combinations of four pearl millet cultivars and four sorghum cultivars. For each combination, all three indices picked out the same species as the dominant one or agreed when there was no dominance. The RCC and LER showed the same pattern of yield advantage or disadvantage, whereas the COA did not. The RCC was not effective in demonstrating the size of the yield advantage, whereas the LER was. Willey (1979) argues that the LER is the most useful index and that it is defined for any set of mixtures and not only replacement series investigations.

Mead and Riley (1981) say that these indices were introduced in the context of a particular crop combination. In intercropping, different crop combinations may need to be compared. An LER may be used for this, but the comparisons may not always be straightforward. Various other indices have been proposed such as crop value and other indices as described in Chapter 13. These are what Kass (1978) calls common parameters.

Example 18.3. The data set of Example 18.1 is used to demonstrate the computation procedure for some of the competition indices, viz.,

(i) the competition ratio CR(i), equation (18.18),
(ii) the coefficient of aggressivity COA(i), equation (18.16), and COA, and
(iii) the relative crowding coefficient K_i = RCC(i), equation (18.6), and RCC, equation (18.7).

Since land equivalent ratios are used in the calculations, they are computed first and are given in the top part of Table 18.3. To illustrate, consider the mixture ABC. $L(h) = L(\mathrm{A}) = 13.88/3(22.58) = 0.205$, $L(i) = L(\mathrm{B}) = 2.08/3(1.64) = 0.423$, $L(j) = L(\mathrm{C}) = 52.54/3(39.54) = 0.443$, and LER(ABC) $= 0.205 + 0.423 + 0.443 = 1.07$. The number 3 in the denominator is used to put the sole crop yield on a 12-plant or sole crop population basis. The three mixtures ABE, ACF, and AEF had the largest LERs. Using the above computed values, the competition ratio or CR(h) values are computed next. For $n = 3$ crops in a mixture with each crop occupying equal area, one-third of the area (4 plants out of 12) is devoted to a cultivar. Therefore, $p_h = \frac{1}{3}$, $1 - p_h = \frac{2}{3}$, and $\mathrm{CR}(h) = L(h)/(\frac{1}{3})/(L(i)/(\frac{1}{3}) + L(j)/(\frac{1}{3}))/2 = 2L(h)/(L(i) + L(j))$. For the mixture ABC, $\mathrm{CR(A)} = 2L(\mathrm{A})/(L(\mathrm{B}) + L(\mathrm{C})) = 2(0.205)/(0.423 + 0.443) = 0.47$, $\mathrm{CR(B)} = 2(0.423)/(0.205 + 0.443) = 1.31$, $\mathrm{CR(C)} = 2(0.443)/(0.205 + 0.423) = 1.41$, and $\mathrm{CR} = 0.47 + 1.31 + 1.41 = 3.19$. In the absence of any competition or cooperation, each $L(h)$ should equal $\frac{1}{3}$, each CR(h) should equal 1, and CR should equal $n = 3$. Mixture ABE had the largest CR value, i.e., 3.29. The four mixtures ABC, ABF, ACE, and AEF all had approximately the same CR values, i.e., 3.2.

A cultivar coefficient of aggressivity from equation (18.16) for a polyculture of $n = 3$ cultivars is obtained as $\mathrm{COA}(h) = L(h)/p_h - (L(i)/p_i + L(j)/p_j)/(n-1 = 2) = 3\{2L(h) - (L(i) + L(j))/2\}$ for equal amounts of space, $\frac{1}{3}$, for each of the three cultivars. Omitting the muliplier 3, the COA(h) values are given in Table 18.3. For the mixture ABC, these are obtained as $\mathrm{COA}(h) = \mathrm{COA(A)} = \{2(0.205) - (0.423 + 0.443)\}/2 = -0.23$, $\mathrm{COA}(i) = \mathrm{COA(B)} = \{2(0.423) - (0.205 + 0.443)\}/2 = 0.10$, and $\mathrm{COA}(j) = \mathrm{COA(C)} = \{2(0.443) - (0.205 + 0.423)\}/2 = 0.13$. Note that the sum of the COA values for a mixture is zero. The largest negative value, -0.23, occurred in the mixture ABC for cultivar A, and the largest positive COA value, 0.29, occurred in the mixture ABE for cultivar B. The largest range of COA values occurred with mixtures ABC, ABE, and AEF.

TABLE 18.3. Competition Indices for the Data in Table 18.1.

	Mixture									
Coefficient	ABC	ABD	ABE	ABF	ACD	ACE	ACF	ADE	ADF	AEF
$L(h)$	0.205	0.410	0.303	0.273	0.233	0.332	0.332	0.312	0.305	0.355
$L(i)$	0.423	0.368	0.618	0.339	0.373	0.389	0.447	0.205	0.255	0.612
$L(j)$	0.443	0.296	0.354	0.159	0.330	0.170	0.541	0.268	0.348	0.375
LER	1.071	1.074	1.275	0.771	0.936	0.881	1.320	0.785	0.908	1.342
CR(h)	0.47	1.23	0.62	1.10	0.66	1.15	0.67	1.32	1.01	0.72
CR(i)	1.31	1.04	1.88	1.57	1.33	1.58	1.02	0.71	0.78	1.68
CR(j)	1.41	0.76	0.77	0.52	1.09	0.48	1.39	1.04	1.24	0.78
CR	3.19	3.03	3.27	3.19	3.08	3.21	3.08	3.07	3.03	3.18
COA(h)	−0.23	0.08	−0.18	0.02	−0.12	0.04	−0.16	0.08	0.00	−0.14
COA(i)	0.10	0.02	0.29	0.12	0.09	0.14	0.01	−0.09	−0.07	0.25
COA(j)	0.13	−0.09	−0.11	−0.15	0.03	−0.19	0.15	0.01	0.07	−0.11
RCC(h)	0.52	1.39	0.87	0.75	0.61	0.95	0.99	0.91	0.88	1.10
RCC(i)	1.47	1.16	3.24	1.03	1.19	1.27	1.62	0.52	0.68	3.15
RCC(j)	1.59	0.84	1.10	0.38	0.99	0.41	2.36	0.73	1.07	1.20
RCC	1.22	1.35	3.10	0.29	0.72	0.49	3.78	0.35	0.64	4.16

Relative crowding coefficients for each cultivar, RCC(h), are computed from equation (18.6). For $n = 3$ and equal space allocated to each cultivar, RCC(h) $= 2L(h)/(1 - L(h))$. For the mixture ABC, RCC(A) $= 2(0.205)/(1 - 0.205) = 0.52$, RCC(B) $= 2(0.423)/(1 - 0.423) = 1.47$, RCC(C) $= 2(0.443)/(1 - 0.443) = 1.59$,

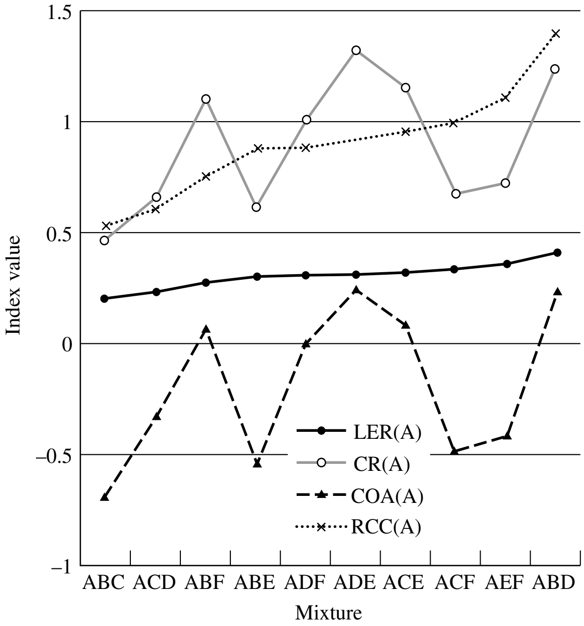

FIGURE 18.6. Values of RCC(A), L(A), CR(A), and COA(A) for the 10 mixtures arranged in ascending order of L(A).

and RCC =0.52(1.47)(1.59) $=$ 1.22. The three mixtures ABE, ACF, and AEF had the highest RCC values and also the largest LER values.

Since cultivar A appeared in all 10 mixtures, a graphical representation of the various statistics L(A), CR(A), COA(A), and RCC(A) from Table 18.3 is given in Figure 18.6. The other cultivars only have four such values and therefore are not included. The COA(A) values have been multiplied by 3 to put them on the same basis as the other competition indices and as given in equation (18.18). The patterns for L(A) and RCC(A) are the same but with larger differences between adjacent mixtures for RCC(A); i.e., the pattern is steeper for RCC(A) than for L(A). The patterns for COA(A) and CR(A) are the same. However, the two pairs of patterns are quite different in the rankings of the mixtures, meaning that they are measuring different characteristics.

Instead of using single cultivar responses in Figure 18.6, the mixture responses for the 10 mixtures are plotted in Figure 18.7 for LER, RCC, and CR. The first two gave the same ranking of the mixtures even though the difference between ranks was much larger for RCC. CR did not rank the mixtures in the same order as the other two measures. The usefulness of the CR measure is not apparent in this example. The individual CR(h) values appear more useful.

In order to obtain an error variance for any of these indices, it is suggested that the index be computed for each cultivar and each mixture in each replicate. Then, an ANOVA may be computed on these values to obtain estimated error variances for

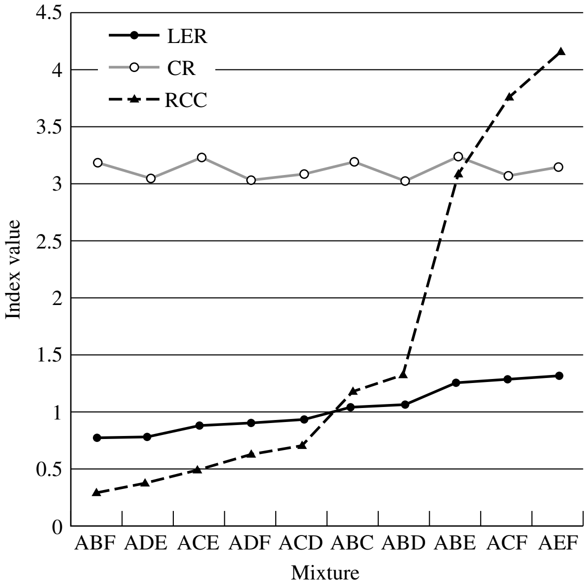

FIGURE 18.7. Values of LER, CR, and RCC for the 10 mixtures in Table 18.1 with mixtures arranged in ascending order of LER.

$L(h)$, CR(h), COA(h), RCC(h), LER, CR, and RCC. The statistical distributions for these quantities are unknown, but since ANOVA procedures are fairly robust as the number of observations increases, this procedure should suffice as a reasonable approximation. Significance tests may be made to determine if CR(h) differs from 1, if $L(h)$ differs from $1/n$, if RCC(h) differs from 1, and if COA(h) differs from 0. $L(h)$ values will become outliers in the sense of becoming very large when the denominators are near zero. The residual mean square becomes inflated and most likely biased upward. $L(h)$ values will be affected adversely when the denominators are near zero.

18.5 Linear Programming

Linear programming had its beginnings in the 1940s (e.g., see Campbell, 1965, Glicksman, 1963, and Manakata, 1979) and has found usefulness in many fields of inquiry. The idea of minimizing cost or material or of maximizing profit or gain from alternative settings has become popular. One use of these ideas (Sprague and Federer, 1951) was to maximize genetic advance by optimally allocating the number of replicates, locations, years, and maize variety entries. An example of using linear programming for intercropping studies is given by DeSilva and Liyange (1978). Paraphrasing Manakata's (1978) description and putting it in terms of intercropping, a linear programming formulation generally has this form:

(i) v decision variables $x_1, x_2, ..., x_v$ all x greater than or equal to zero, (hectares of each of v cultivars) are being considered.
(ii) v equations or constraints relating the $x_1, x_2, ..., x_v$ to the desired amounts $r_1, r_2, ..., r_p$ of the p products; p may be equal, less than, or greater than v.
(iii) An objective function is available to maximize or minimize.

For example, a farmer may wish to grow $v = 3$ crops, maize (M), beans (B), and castrol beans (C) and he desires to produce 10,000 kg of starch, 4000 kg of protein, and 2000 kg of oil. Suppose M has 60% starch, 4% oil, and 10% protein, C has 50% oil, 5% protein, and 10% starch, and B has 50% protein, 10% oil, and 20% starch (see, e.g., Martin and Leonard, 1949). From previous year's yield data or from experiments such as the one in Example 18.1, the kilograms per hectare of starch, protein, and oil can be determined for each of the three crops. Armed with such information, the farmer is in a position to determine how to optimally allocate his hectares to achieve the desired goals. This type of information is useful for government officials desiring to have specified amounts of p products given that v cultivars are to be grown in their region. A cost function, H, for growing each of the v cultivars is 1 ha per cultivar. The problem may be set up as follows:

	Cultivars					
Product	1	2	3	$\cdots$	v	Goal = **R**
1	a_{11}	a_{12}	a_{13}		a_{1v}	r_1
2	a_{21}	a_{22}	a_{23}		a_{2v}	r_3
3	a_{31}	a_{32}	a_{33}		a_{3v}	r_3
				$\cdots$		
p	a_{p1}	a_{p2}	a_{p3}		a_{pv}	r_p
Cost = **H**	1	1	1		1	

A is a p-by-v matrix of coefficients a_{ij}, **H** is a row vector of ones, **R** is a column vector of the r_i, and **X** is a column vector of the unknowns x_j. The linear programming problem is to minimize **HX**, the objective function, subject to the restrictions $\mathbf{AX} \geq \mathbf{R}$ and **X**, and $\mathbf{AX} \geq \mathbf{R}$ represents the set of equations. The *simplex method* (e.g., Campbell, 1965, Glicksman, 1963, and Manakata, 1979) was devised to obtain solutions for such linear programming problems and is described below. The first step is to set up the problem, which for our case is the matrix

$$\begin{pmatrix} \mathbf{A}' & \mathbf{I}_p & \mathbf{H}' \\ \mathbf{R}' & \mathbf{0} & f \end{pmatrix},$$

where $\mathbf{A}'$ is the transpose of the matrix **A**, $\mathbf{I}_p$ is the identity matrix (ones on the diagonal and zeros elsewhere), **0** is a row vector of zeros, and f is a scalar which for our case turns out to be the unknown total area in hectares. To solve for the **X** matrix, the following steps are used:

1. Choose any column of $\mathbf{A}'$ whose last entry is positive.
2. Find a *pivot* entry by dividing each nonzero entry of the selected column into the entry in $\mathbf{H}'$ and selecting the entry which has the smallest non-negative

value. The row which includes the pivot is the *pivot row* and its values remain. The entries in the last column of $\mathbf{H}'$ need to be non-negative.
3. Make all entries in the selected column other than the pivot equal to zero. This is done by a matrix row operation on the nonpivot rows.
4. Repeat steps 1, 2, and 3 on all other columns involving $\mathbf{A}'$ until all entries in the last row are nonpositive, i.e., less than or equal zero.

The solutions for $\mathbf{X}$ will appear in the part formerly occupied by the $\mathbf{0}$ vector. Setting the result formerly occupied by f equal to zero, a solution for f, the total effective area (TEA), a sum of the entries of $\mathbf{X}$, is obtained. When $v = p$, we may obtain a solution for $\mathbf{X}$ as $\mathbf{X} = \mathbf{A}^{-1}\,\mathbf{R}$. However, the simplex procedure is described first using the following example.

Example 18.4. Suppose three sole crops castrol beans (C), maize (M), and beans (B) are to be grown and it is desired to produce 10,000 kg of starch, 4000 kg of protein, and 2000 kg of oil. Further suppose that the following amounts are produced:

Product	C	M	B	Goal
Starch	180	1,300	200	10,000
Protein	90	200	500	4,000
Oil	900	80	100	1,000
Cost = hectares	1	1	1	

Putting the above in matrix form, we have

$$\mathbf{A} = \begin{pmatrix} 180 & 1300 & 200 \\ 90 & 200 & 500 \\ 900 & 80 & 100 \end{pmatrix}, \quad \mathbf{R} = \begin{pmatrix} 10{,}000 \\ 4,000 \\ 2,000 \end{pmatrix}, \quad \text{and} \quad \mathbf{H}' = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Following the general procedure outlined above, we have

Step 1

$$\begin{pmatrix} \mathbf{A}' & \mathbf{I}_p & \mathbf{H}' \\ \mathbf{R}' & \mathbf{0} & f \end{pmatrix} = \begin{pmatrix} 180 & 90 & 900 & 1 & 0 & 0 & 1 \\ 1,300^{*} & 200 & 80 & 0 & 1 & 0 & 1 \\ 200 & 500 & 100 & 0 & 0 & 1 & 1 \\ 10{,}000 & 4,000 & 2,000 & 0 & 0 & 0 & f \end{pmatrix}$$

Step 2 Select column one. Then 1/1300 is less than 1/180 and 1/200, making 1,300 the pivot and row two the pivot row. An asterisk is used to denote this.

Step 3 Make the entries of the selected column other than 1300 equal to zero by subtracting the appropriate multiple of row 2 ($R2$) from each of the other rows (Ri) as follows:

$$R1 - 180R2/1300 = R1 - 9R2/65,$$

$$R3 - 200R2/1300 = R3 - 2R2/13,$$

$$R4 - 10{,}000R2/1300 = R4 - 100R2/13.$$

Using the above matrix row manipulations, we obtain,

$$\begin{pmatrix} 0 & 62.308 & 888.923 & 1 & -0.1385 & 0 & 0.8615 \\ 1,300 & 200 & 80 & 0 & 1 & 0 & 1 \\ 0 & 469.231^{*} & 87.692 & 0 & 0.1538 & 1 & 0.8462 \\ 0 & 2,461.538 & 1,384.615 & 0 & -7.6923 & 0 & f-7.6923 \end{pmatrix}.$$

Step 4 Select column two and find the pivot which is 469.231. Make all entries in column zero except 469.231 by the following row operations:

$$R1 - 62.308R3/469.231,$$

$$R2 - 200R3/469.231, \text{ and}$$

$$R4 - 2,461.538R4/469.231.$$

From these row operations, the following results:

$$\begin{pmatrix} 0 & 0 & 877.279^{*} & 1 & -0.1181 & -0.1328 & 0.7491 \\ 1,300 & 0 & 42.623 & 0 & 1.0656 & -0.4262 & 0.6402 \\ 0 & 469.231 & 87.692 & 0 & -0.1538 & 1 & 0.8462 \\ 0 & 0 & 924.592 & 0 & -6.8855 & -5.2459 & f-12.1314 \end{pmatrix}.$$

Step 5 Select a pivot in column three which is 877.279. Perform the following row operations

$$R2 - 42.623R1/877.279,$$

$$R3 - 87.692R1/877.279, \text{ and}$$

$$R4 - 924.592R1/877.279.$$

to obtain

$$\begin{pmatrix} 0 & 0 & 877.279 & 1 & -0.1181 & -0.1328 & 0.7491 \\ 1,300 & 0 & 0 & -0.0486 & 1.0713 & -0.4197 & 0.6038 \\ 0 & 469.230 & 0 & -0.1000 & -0.1420 & 1.0133 & 0.7713 \\ 0 & 0 & 0 & -1.0539 & -6.7610 & -5.1039 & f-12.9209 \end{pmatrix}.$$

The number of hectares of C is 1.0539, of M is 6.7610, and of B is 5.1039. Their sum is 12.9209 or approximately 13 hectares. The TEA is 12.9209 hectares.

When the number of crops equals the number of products, i.e., $v = p$, the solution for the following three equations results in the above solutions within rounding error:

$$180x_1 + 1300x_2 + 200x_3 = 10{,}000,$$

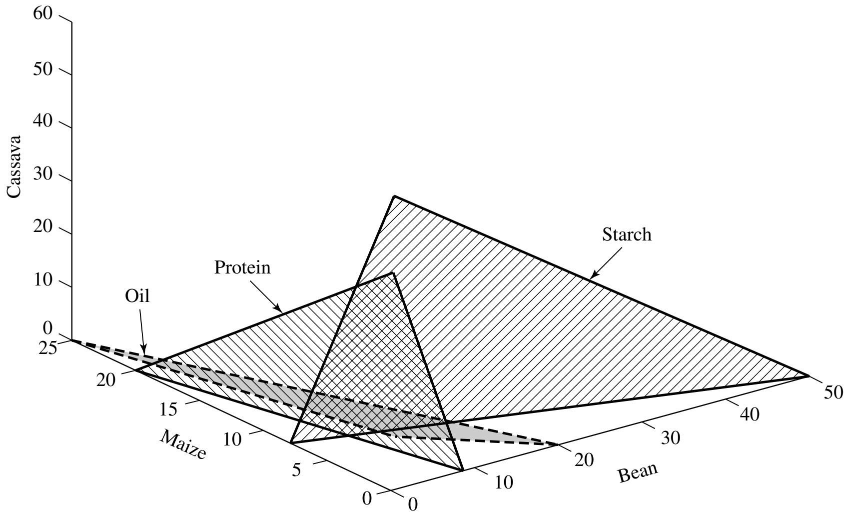

FIGURE 18.8. Hectares of C, M, and B to obtain 10,000 kg of starch, 4000 kg of protein, and 2000 kg of oil.

$$0x_1 + 200x_2 + 500x_3 = 4000,$$

$$900x_1 + 80x_2 + 100x_3 = 2000.$$

To obtain solutions for these three equations, the following GAUSS program was used:

Let A[3,3] = 180 1,300 200 90 200 500 900 80 100;
Let R[3,1] = 10,000 4,000 2,000; **X** = inv(**A**)*$\mathbf{R}$; **X**′;

The solution obtained was $\mathbf{X}' = [1.0539 \quad 6.7608 \quad 5.1060]$, which is equal to the above solutions within rounding error. For the three crops B, M, and C, a TEA may be obtained graphically. A graphical solution is the intersection of the three planes as drawn in Figure 18.8.

Example 18.5. For the particular three crops in Example 18.1, the protein and starch from castrol beans is not suitable for human or animal consumption (Martin and Leonard, 1949). Also, the oil in maize and bean when used for human consumption is not recoverable. Consider the cropping system wherein the mixture C:M:2B and the two sole crops M and B are to be grown. The object is to reach the desired goal of producing 10,000 kg of starch, 4000 kg of protein, and 2000 kg of oil under the above restrictions. The problem is to determine the hectarage for each crop such that a minimum number of hectares will be used and to determine if this number is less than growing the three sole crops as in Example 18.4. Suppose the data are as follows:

	Sole		C:M:2B			
Product	M	B	M	B	C	Total
Starch	1,300	200	800	100	0	900
Protein	200	500	150	250	0	400
Oil	0	0	0	0	600	600

The solution for number of hectares for each crop is obtained by solving the following three equations:

$$1300x_1 + 200x_2 + 900x_3 = 10{,}000,$$

$$200x_1 + 500x_2 + 400x_3 = 4000,$$

$$600x_3 = 2000.$$

The solution for x_3 is 2000/600 = 10/3 = 3.3333 hectares of the mixture in order to reach the desired goal of 2000 kg of oil. $x_1 = 4.8634$ hectares of maize and $x_2 = 3.3879$ hectares of beans and 3.3333 hectares of the mixture are required to reach the goal of 10,000 kg of starch and 4000 kg of protein. The total number of hectares required is TEA = 4.8634 + 3.3879 + 3.3333 = 11.5847, or 11.6 hectares. This is a savings of 1.3 hectares, or 10%, over the cropping system of sole crops in Example 18.4.

The above comparison does not take into account the starch and protein produced by C and the oil produced by M and B as was done to find the TEA in Example 18.4. A more appropriate comparison is to find the TEA for the following data set:

	Sole		C:M:2B			
Product	M	B	C	M	B	Total
Starch	1300	200	120	800	100	1020
Protein	200	500	60	150	250	460
Oil	80	100	600	60	50	710

The three equations to be solved are

$$1300x_1 + 20x_2 + 1020x_3 = 10{,}000,$$

$$200x_1 + 500x_2 + 460x_3 = 4000,$$

$$80x_1 + 100x_2 + 710x_3 = 2000.$$

The solutions are $x_1 = 5.8125$ hectares of maize, $x_2 = 4.2347$ hectares of beans, and $x_3 = 1.5655$ hectares of the mixture C:M:2B. The TEA $= 5.8125 + 4.2347 + 1.5655 = 11.6127$, or 11.6 hectares which is a 10% savings over using only the three sole crops. The result did not change over what was obtained above for this particular example.

18.6 Discussion

The yield results of Example 18.2 are striking, indicating that these particular three cultivars demonstrated large cooperative effects. The mixtures of two cultivars do not indicate the results for mixtures of three crops, again demonstrating that inferences from polycultures of two cultivars may not be useful for making inferences of results for polycultures of three or more cultivars. The beneficial effects from intercropping with the appropriate polycultures can be considerable as indicated by this example. When a treatment design contains some or all possible combinations of the m cultivars under consideration, general mixing ability, bi-specific mixing, etc. effects may be obtained. When few combinations such as those described in Example 18.2 are available, the methods of this chapter will be useful.

Federer and Basford (1991) and the references cited therein consider competition effects in a different manner than in this chapter. This means that the subject of competition has many aspects and more than considered there and herein exist. There are the problems of how competition enters and is exhibited and what are the factors contributing to competition. The effects of intimacy and spatial arrangements need to be considered in combination with competition effects. If competition effects are inhibitive, spatial arrangements may be used to eliminate them. If competition effects are cooperative, spatial arrangements and intimacy may be used to enhance them.

The calculations involved in the simplex method may appear tedious, but with the aid of computers and computer software, these calculations may become routine. Linear programming, even at the farm level, can be very beneficial to subsistence farmers who have limited hectarage and who need sufficient food for the family. Optimum allocations of sole and mixture crops may be computed for a number of scenarios and distributed to farmers for guidance in allocating their limited resources.

18.7 Problems

18.1 Using the data for cultivar E in Table 13.3, obtain the same calculations as were obtained for cultivar A in Example 18.1 and compare the results obtained for the two cultivars A and E. Give a graphical presentation of results.

18.2 Complete the calculations described in Example 18.2 for *each* of the three replicates. Then, obtain an ANOVA for these computed values and construct standard errors of differences. Which, if any, of the mixture yields are significantly different from predicted at the 10% level and at the 5% level of significance? Give a graphical presentation of the results.

18.3 For the data of Example 18.2, compare the 12 entries with regard to the mean method, the generalized half-hectare method, and the pure stand production method. How do they agree with the LER method? Give a graphical presentation of the results.

18.4 Replace the goal of 10,000 kg of starch, 4000 kg of protein, and 2000 kg of oil with 10,000 kg of starch, 2000 kg of protein, and 5000 kg of oil.

Perform the calculations described in Example 18.4 for the new goals. Give a graphical presentation of the results.

18.5 Using the revised goal described in Problem 18.4, do the calculations described in Example 18.5. Give a graphical presentation of the results.

18.6 For the mixtures ABD and BDF, the following yields are available:

Cultivar	Mixture	Sole	Cultivar	Mixture	Sole
A	27.75	22.58	B	1.62	1.64
B	1.81	1.64	D	0.47	2.71
D	2.41	2.71	F	0.44	1.59

The yields are on a four-plant basis. The sole yields need to be multiplied by 3 to obtain a 12-plant or sole crop basis. Prepare graphs similar to Figures 18.1 and 18.3 showing cooperation or inhibition.

18.7 Consider the following two mixture and sole crop yields:

Cultivar	Mixture	Sole	Cultivar	Mixture	Sole
B	4.00	1.64	B	1.64	1.62
C	60.75	39.54	D	2.71	0.47
E	9.19	8.96	F	1.59	0.44
Total	73.94	50.14		5.94	2.53

Prepare a graph similar to Figure 18.5 depicting the mixture yields and their predicted sole yields and discuss any possible cooperative, inhibitive, neutral, or compensatory effects.

18.8 Literature Cited

Atkinson, A. C. (1996). The usefulness of optimal experimental designs. *J. Roy. Statist. Soc., Series B* **58**, 59-76

Aitkinson, A. C. and A. N. Donev (1992). *Optimal Experimental Designs*. Oxford University Press, Oxford.

Campbell, H. G. (1965). *An Introduction to Matrix Vectors and Linear Programming*. The Appleton-Century Mathematics Series, Appleton-Century-Crofts, New York.

Cornell, J. (1990). *Experiments With Mixtures: Designs, Models, and Analysis of Mixture Data*, second edition, John Wiley & Sons, Inc., New York.

DeSilva, N. T. M. H. and M. de S. Liyange (1978). An appraisal of linear programming in planning intercropping. *Agric. Agroid J.* **11**(4), 3-7.

de Wit, C. T. (1960). On competition. *Versl. Landbouw. Onderzoekingen, Wageningen* **66**, 1-82

de Wit, C. T. (1961). Space relationships within populations of one or more species. In *Symp. Soc. for Exp. Biol. XV. Mechanisims in Biological Competition*. Academic Press, New York, pp. 314-329.

de Wit, C. T. and J. P. van den Bergh (1965). Competition among herbage plants. *Netherlands J. Agric. Sci.* **13**, 212-221.

Donald, C. M. (1963). Competition among crop and pasture plants. *Adv. Agron.* **15**, 1-118.

Federer, W.T. and K.E. Basford (1991). Competing effects designs and models for two-dimensional field arrangements. *Biometrics* **47**, 1461-1472.

Glicksman, A. M. (1963). *Linear Programming and the Theory of Games*. John Wiley & Sons, Inc., New York.

Haizel, K. A. (1974). The agronomic significance of mixed cropping. *Ghana J. Agric. Sci.* **7**, 169-178.

Hall, R. L. (1974a). Analysis of the nature of interference between plants of different species. 1. Concepts and extension of the de Wit analysis to examine effects. *Australian J. Agric. Res.* **25**, 739-747.

Hall, R. L. (1974b). Analysis of the nature of interference between plants of different species. 2. Nutrient relations in a Nandi *Setaria* and Greenleaf *Desmodium* association with particular reference to potassium. *Australian J. Agric. Res.* **25**, 749-756.

IRRI (1974). Multiple cropping. In IRRI Annual Report for 1973, pp. 15-34.

Kass, D. C. L. (1978). Polyculture cropping systems: Review and analysis. *Cornell Int. Agric. Bull.* 32, N.Y. State College of Agriculture and Life Science, Ithaca, NY.

Lipman, J. G. (1912). The associative growth of legumes and non-legumes. *New Jersey Agric. Expt. Station Bull.* **252**, 1-48.

Manakata, T. (1979). *Matrices and Linear Programming with Applications*. Holden-Day, Inc., San Francisco.

Martin, J.H. and W.H. Leonard (1949). *Principles of Field Crop Production*. The MacMillan Company, New York.

McGilchrist, C. A. and B. R. Trenbath (1971). A revised analysis of plant competition experiments. *Biometrics* **27**, 659-671.

Mead, R. and J. Riley (1981). A review of statistical ideas relevant to intercropping research (with discussion). *J. Roy. Statist. Soc., Series A* **144**, 462-509.

Morrish, R. H. (1934). Crop mixture trials in Michigan. *Michigan State College Agric. Expt. Station Bull.* **256**, 1-11.

Niqueux, M. (1959). Choic de varietes d'Arachides au Tchad III. Essais de culture associee d'arachides er de sorghos au Tchad. *L'Agronomie Tropicale* **14**, 501-502.

Pilz, F. (1911). Leguminosen und graminen in Rein- und in Mengsaaten mit besonderer Berucksichtunge der Stickstoffnutzung. *Ztschr. Landw. Versuchs. Osterr*. **14**, 1140-1210.

Soria, J., R. Bazan, A. M. Pinchinat, G. Paez, N. Mateo, R. Moreno, J. Fargas, and W. Forsythe (1975). Investigacion sobre sistemas de produccion agricola para el pequeno agricultor del tropico. *Turrialba* **25**, 283-293.

Sprague, G. F. and W. T. Federer (1951). A comparison of variance components in corn yield trials: II. Error, year × variety, location × variety, and variety components. *Agron. J.* **43**, 535-541.

Van den Bergh, J. P. (1968). An analysis of yields of grasses in mixed and pure stands. *Versl. Landbouw. Onderzoekingen, Wageningen*. **71**, 1-71.

Willey, R. W. (1979). Intercropping—its importance and research needs. Part I. Competition and yield advantages. Part 2. Agronomy and research approaches. *Field Crops Abstr.* **32**, 1-10 and 73-85.

Willey, R. W. and D. S. O. Osiru (1972). Studies on mixtures of maize and beans (*P. vulgaris*) with particular reference to plant populations. *J. Agric. Sci., Cambridge* **79**, 517-529.

Willey, R. W. and M. R. Rao (1980). A competitive ratio for quantifying competition between intercrops. *Exper. Agric*. **16**, 117-125

CHAPTER 19

Application of Intercropping Procedures to Other Fields

19.1 Introduction

The statistical designs and methods utilized in Volume I and in the preceding chapters may be used in many other subject matter fields either directly or with slight modifications. These techniques deal with agricultural systems but are general and may be applied to many other systems. Dealing with the single components of a system one at a time may be inefficient data analysis and may not yield the necessary information. As has been demonstrated in the examples presented, one needs to consider the entire system as a unit in addition to considering individual components of the system. Considering only the individual components of a system is useful but does not ferret out all the necessary information contained in the systems and mixtures under consideration. For example, consider three prescribed drugs, each of which may have no side effects or mixtures of two which may have no side effect, but when all three are combined, the combination becomes carcinogenic. The study of combining abilities (interaction) is essential in some experiments.

Some of the areas using mixtures and systems of treatments are discussed in the following sections. Specifically, the areas discussed are as follows:

(i) agriculture,
(ii) ecology,
(iii) golf courses,
(iv) medicine and pharmacology,
(v) education,
(vi) exercise and aerobics,
(vii) engineering,

(viii) marketing,
(ix) nutrition and diet,
(x) others.

19.2 Agriculture

There are many agricultural systems which involve multiple cropping in one form or another. Some of these are as follows:

(i) rotation cropping,
(ii) relay or sequential cropping,
(iii) intercropping of crops with different lengths of growing seasons,
(iv) pasture studies of grass and grass–legume mixtures,
(v) intercropping in tree (orchard, paper, lumber) plantations,
(vi) double cropping,
(vii) cover crops with main crop or intercrop,
(viii) weeds and crops,
(ix) other areas.

As seen from the above list, agriculture is rife with examples involving systems of multiple cropping situations.

Rotation cropping systems are of many and diverse types; see, e.g., Cochran (1939), Crowther and Cochran (1942), Yates (1949), Patterson (1953, 1959, 1964). To illustrate a type of rotation experiment, consider that three cultivars are to be used and that these are M = maize, S = soybean, and W = wheat. Some possible three-sequence experiment designs are

	Design 1			Design 2			Design 3			Design 4			Design 5		
Yr/Seq.	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
1	M	W	S	M	W	S	M	W	S	M	W	S	M	W	S
2	W	S	M	S	M	W	W	S	M	W	S	M	M	W	S
3	S	M	W	W	S	M	S	M	W	S	M	W	W	S	M
4	M	W	S	M	W	S	M	W	S	S	M	W	W	S	M
5	W	S	M	S	M	W	S	M	W	W	S	M	S	M	W
6	S	M	W	W	S	M	W	S	M	M	W	S	S	M	W
7	M	W	S	M	W	S	M	W	S	M	W	S	M	W	S
8	W	S	M	S	M	W	W	S	M	W	S	M	M	W	S
9	S	M	W	W	S	M	S	M	W	S	M	W	W	S	M
10	M	W	S	M	W	S	M	W	S	M	W	S	M	W	S
etc.															

Design 1 has M following the legume S for the situation where M is the crop of interest. Design 2 has W following the legume S in the cropping sequence but may not be practical if the wheat variety is susceptible to lodging under higher nitrogen levels. These two designs are the most commonly used ones in crop rotation studies. Sometimes, three additional plots will be added to study the effect of continuous cropping (sole cropping) of the three crops. Such continuous cropping plots could be added to any or all of the above five designs. Design 3 was described by Federer and Atkinson (1964) as a tied-double-changeover design. It is of the Cochran *et al.* (1941) and Williams (1949) type of double-changeover design which would use six sequences and place the treatment sequences from years 4, 5, and 6 in sequences 4, 5, and 6. Direct effects and carryover effects of each of the three crops are obtainable with the double-changeover designs. Designs 3, 4, and 5 allow the estimation of carryover effects. Federer and Kershner (1998) show that direct-by-residual and cumulative treatment-by-period interactions are obtainable when using Designs 4 or 5. More sequences of crops as described for the designs of Federer and Kershner (1998) may be used to measure direct, residual, and cumulative effects in a shorter time period of years. Note that the effects and interactions pertain to the particular genotypes included in the treatment design for the rotation. Also, note that several rotational systems may be included in the experiment by adding more sequences of crops, with comparisons between systems being the ones of interest. A base or standard rotation may be used for comparison of rotational systems in place of the continuous cropping system.

A land equivalent ratio (LER) is useful in summarizing the results from rotation experiments when continuous cropping sequences are included in the experiment design. Let Y_{ir} be the yield of crop i, i = M, S, W, in the rotation and let Y_{ic} be the yield of crop i when grown continuously on the same land area. If rotation cropping is successful, the ratio $R_i = Y_{ir}/Y_{ic}$ for crop i should be greater than 1. Since there will be ratios R_{hi} for each year h, the ratios R_{hi} may be plotted graphically against year h and a regression of the ratios R_{hi} on year obtained. A form of asymptotic regression would be more appropriate than polynomial regression. The interest is in the intercept, slope, curvature, and limiting value (asymptote) of the ratios R_{hi} over time. To evaluate the cropping system, a sum of the ratios R_{hi} over crops would be used for each year's LER and a regression of the yearly LERs on years could be obtained. The estimated direct, residual, and cumulative effects of the crops could also be used in the LER analyses described above. For Design i, the carryover or residual effects from S to M would be expected to be positive, as there is a yield benefit of growing a legume preceding the growing of a grass species. The carryover effect from M to W would likely be negative, as would the carryover effect from W to S. If several rotational systems are included in the experiment, comparisons among the intercepts, slopes, curvatures, and limiting values of each of the systems may be made.

In addition to yields, the experimenter may be interested in the economic or value returns of the system. Let p_{hi} be the value (or relative value to a base crop) of crop i in year h. Then, for crop i in year h, the economic gain is $V_{hi} = p_{hi}(Y_{hir} - Y_{hic})$. It may be desirable to standardize values by using ratios of prices as was explained in Chapter 13 and in Chapter 4 of Volume I. The sum of the V_{hi} over i gives the yearly values, $V_{h\cdot}$, of the system. Regressions of the $V_{h\cdot}$ or the V_{hi} on years may be used to further summarize the results. The ordinary t-statistic may be used to compare crop i yields or values in a rotation and in continuous cropping or between the yields or values of a crop in two different rotations.

Benefits other than increased yields and values are associated with crop rotations. Better soil structure, lower soil erosion, and weed control are three possible benefits. The increase in nutritional value of a grass following a legume over continuous cropping is another possibility. The trait soil structure may be hard to quantify, but perhaps some measure such as the organic matter content of the soil could be used as a measure of soil structure. The value of weed control, soil structure, and soil erosion may also be quantified through crop yield. In addition to crop rotational systems of intercropping, several other agricultural practices involve the sequential growing of the same or different cultivars in the same or different years. One such system is *double cropping*, which involves the growing of two crops sequentially in the same year. Examples are wheat–wheat, soybean–soybean, soybean–wheat, etc. A possible experiment design to study double cropping for two crops such as W = wheat and S = soybean is as follows:

	Sequence									
Year	1	2	3	4	5	6	7	8	9	10
1	W	W	S	S	S	W	S	W	W	S
	W	—	S	—	W	S	—	—	W	S
2	W	W	S	S	S	W	W	S	S	W
	W	—	S	—	W	S	—	—	S	W
3	W	W	S	S	S	W	S	W	W	S
	W	—	S	—	W	S	—	—	W	S
4	W	W	S	S	S	W	W	S	S	W
	W	—	S	—	W	S	—	—	S	W
5	W	W	S	S	S	W	S	W	W	S
	W	—	S	—	W	S	—	—	W	S
etc.										

For the design in sequences 1 and 3, sequences 2 and 4 could be considered as check treatments for continuous cropping by standard methods, and sequences 7 and 8 could be check treatments for a two-crop rotational system for the standard method. Sequences 5 and 6 represent half-year rotational systems The other sequences involve double cropping procedures. Sequences 9 and 10 involve double cropping the same crop within a year in a rotational system over years. Note that the growing season for sequences 2, 4, 6, and 8 may overlap those in the other sequences, as the optimal growing season for the one crop per year would be selected from the entire year and not limited to either the first half or the second half of the year. Various combinations of the above sequences may be used for the treatment design, depending on the goals of the experimenter. The above ideas are directly extendible to three, four, or more cultivars in an agricultural system.

Ideas of summarizing the results from previous chapters may be used here as well. Let Y_{hij} equal the response (yield) for crop h in year i and sequence j. Land use evaluations using an LER for year i would be of the form Y_{Wi1}/Y_{Wi2} and Y_{Si3}/Y_{Si4} for continuous cropping, where Y_{Wi1} is the sum of the two wheat crops in year i, Y_{Si3} is the sum of the two soybean crops in year i, Y_{Wi2} is the yield of W from one crop in year i, and Y_{Si4} is the yield of S from one crop in year i. For double cropping to be beneficial, each of the ratios should be greater than unity and by a large enough margin to more than recover the additional costs of production of double cropping over single cropping. For the rotational cropping system, the ratio of the W yields in sequences 5 and 6 (or 9 and 10) to those in sequences 7 and 8 form the LER. The same form is also used for the S yields. LERs of rotational cropping to continuous cropping may be obtained using sequences 2 and 4 yields for the denominator of the LER. For sequences 5 to 10, an LER for the system may be formed from the sum of the ratios for the two crops S and W.

Economic, nutritional, or other evaluations of double cropping relative to single cropping may be made as well. Wheat grown in a continuous double cropping system may be lower in protein content than when grown in a single cropping system or in a rotational system. The cost of producing a double crop is higher than for one crop. Let p_{hij} be the profit from crop h in year i and sequence j. Then, $p_{hij}Y_{hij}$ is the profit derived from h in year i and sequence j. Ratios and differences of the $p_{hij}Y_{hij}$ values from the appropriate sequences may be used to summarize the results from an experiment. Values other than profit may be evaluated in a similar manner. In addition to the above analyses, regressions of the above statistics on year may be used as described previously to obtain the intercept, slope, curvature, and limiting values of responses.

Sequential cropping sytems using a cover crop is another topic receiving considerable attention in the literature (see Chapter 20). Cover crops are grown in the off-season for the crop of interest. They are used to control soil erosion, to improve the soil structure, and to provide nutrients for the following crop. For example, a legume may be planted in the fall after a maize crop is harvested, it may be used for pasture or not, and the residue used as a green manure crop for the following crop of maize or oat, e.g., which is spring planted. Also, instead of plowing the cover crop under, a form of no-till or minimum tillage may be used in connection with the cover crop. Some cover or green manure cropping systems that may be useful for selecting a treatment design for three cultivars M = maize, W = wheat, and S = soybean are as follows:

	Sequence											
Year	1	2	3	4	5	6	7	8	9	10	11	12
1	C	C	C	—	—	—	C	C	C	—	—	—
	M	S	W	M	S	W	M	S	W	M	S	W
2	C	C	C	—	—	—	C	C	C	—	—	—
	M	S	W	M	S	W	W	M	S	W	M	S
3	C	C	C	—	—	—	C	C	C	—	—	—
	M	S	W	M	S	W	S	W	M	S	W	M
4	C	C	C	—	—	—	C	C	C	—	—	—
	M	S	W	M	S	W	M	S	W	M	S	W
5	C	C	C	—	—	—	C	C	C	—	—	—
	M	S	W	M	S	W	W	M	S	W	M	S
etc.												

where C is the cover crop. Sequences 1–6 involve continuous cropping of M, S, and W. Sequences 7–12 are three crop rotations with sequences 7–9 including a cover crop and sequences 10–12 without a cover crop. Sequences 1–3 and 7–9 may be duplicated and no-till versus conventional tillage, fertilizer versus no fertilizer, weed control versus minimum weed control, etc. treatments included in the experiment.

For land use, protein content, calorie content, soil structure, soil erosion, etc., the methods described above are directly applicable. If the economic value of a cover crop system is used to summarize the result, the cost of growing a cover crop, the value of the cover crop, and any other cost or value from the cover crop need to be taken into account. Let p_{hij} be the profit from crop h in year i from sequence j as before. Then, differences $p_{hij}Y_{hij} - p_{hij'}Y_{hij'}$, $j \neq j'$, or ratios $p_{hij}Y_{hij}/p_{hij'}Y_{hij'}$ may be used to summarize the results for h = M from sequences 1 and 4, for h = S from sequences 2 and 5, and for h = W from sequences 3 and 6. The M, S, and W responses from sequences 7–9 may be compared with those from sequences 10–12 to obtain the value of using cover or green manure crops in a crop rotation system. In addition, to individual crop comparisons, the total profit from all crops may be used in comparisons for sequences 7–12. The latest year's results or limiting value estimates for each crop as well as for all crops are useful summary statistics for a cropping system.

Weeds (plants out of place) form a naturally occurring intercropping system. Many studies are conducted on various forms of weed control or weed elimination. The less that is spent on weed control, the greater the return from a crop. This has led to studies on minimum weed control to determine the level and density of weeds which do not detract from cultivar response. Just as in other intercropping systems, certain levels and types of weeds may enhance crop responses. If the presence of certain types of weeds is related to disease or insect infestation and/or to erosion control, this is beneficial for the crop of interest. The economics and effects of minimum weed control systems need to be evaluated.

Another area of intercropping which receives considerable attention in the literature is pasture studies (see Chapter 20). Mixtures of several grasses and/or several legumes in various combinations may form the treatment design of a pasture experiment. In addition to mixture content, management, fertilizer, and other variables may be investigated. Some of the treatment designs from Chapters 15 and 16 may be useful for pasture studies. The concepts of general and specific mixing abilities are useful here, just as they were for the experiments described in previous chapters. The specific goals of a study determine the appropriate treatment design for the investigation.

There are several ways in which a particular pasture treatment response may be measured. Some of these are as follows:

(i) weight of harvested forage for the pasture mixture,
(ii) weights of the individual grasses or legumes in the mixture,
(iii) carrying capacity (number of animals) of the pasture mixture,
(iv) pounds of beef, mutton, etc. produced for the pasture mixture,
(v) nutritional content of the pasture mixture,

(vi) economic return from a pasture mixture,
(vii) erosion control for a pasture mixture,
(viii) other.

Except for item (ii), only the total for a mixture is available, as it was for the designs and analyses of Chapter 16. If sole crop responses for the cultivars in a mixture are available, the mean method of comparing the mixture with sole crop yields is a candidate for summarizing the results. For mixtures of two and three crops, a mean value statistic (Chapter 18) is obtained as $2Y_{12}/(Y_{1S}+Y_{2S})$ and $3Y_{123}/(Y_{1S}+Y_{2S}+Y_{3S})$, respectively, where Y_{12} and Y_{123} are responses for mixtures of two and three, respectively, and Y_{hS}, for $h = 1, 2, 3$, is the sole crop h response. Alternatively, the differences $(2Y_{12} - Y_{1S} - Y_{2S})$ and $(3Y_{123} - Y_{1S} - Y_{2S} - Y_{3S})$ may be used to make comparisons. If the response is weight of produce, taking the mean weight of the sole crops may not be meaningful (Kass, 1978). If the cultivars have similar weights, then taking an average of sole crop yields may be meaningful. Since this is true for some pasture mixtures, the mean method may provide a useful tool for summmarizing the results. Even for item (ii) above, hand separation of the components of a hay mixture can be tedious, time-consuming, and costly, and the experimenter may elect to use the total rather than individual component weights for the mixture. In such cases, as well as for the other responses listed above, the above ratios and differences should suffice for summarization of results. These sums and differences may be computed each year and regressions performed as described above.

Another important intercropping system is the interplanting of cultivars in orchards for fruit, forests for lumber, trees for paper, tea plantations, coffee plantations, grape vineyards, etc. (see Chapter 20). The intercropped cultivars may or may not be harvested, pastured, or used in other ways. Instead, the intercrops may be used to control erosion, for insect control, for disease control, as a shade or supporting mechanism, or for other purposes.

Relay cropping (the interplanting of a second or even third crop into the area where the first crop is growing) is another method of intercropping. For example, in a maize–bean intercrop, where the beans are harvested before maize and with 1.75 meters, say, between pairs of maize rows, a third crop like oats may be planted before the maize is harvested but after beans are harvested. In garden plantings, relay cropping of a number of cultivars may be practiced with the idea of having a continuous production of vegetables.

Some intercropping systems involve cultivars with different growing seasons. For example, sugarcane may have a 1-year or a 2-year season. Cassava has a 1-year season. These crops may be intercropped with crops which require a 3-month, a 4-month, or even a 6-month growing season. Since a crop like sugarcane can take 4–5 months to "close in" (develop a canopy), there is ample time to include one or two short-season crops. Garden peas followed by onions may be used in sugarcane plantings. Potatoes, dry beans, soybeans, maize for roasting ears, etc. are crops that may be intercropped with sugarcane, cassava, and on rice paddies. When one crop's growing season is twice as long as a second crop and when two crops of the second crop is grown, land use, economic value, nutritional, and other values are still useful for summarizing the results. For example, suppose two crops of beans (melon, or maize) are intercropped with cassava. A land use statistic would be

$$\text{LER} = L_c + L_b = Y_{cm}/Y_{cs} + (Y_{b1m} + Y_{b2m})/(Y_{b1s} + Y_{b2s}),$$

where Y_{cm} and Y_{cs} are cassava yields in the mixture and as a sole crop, respectively; Y_{b1m} and Y_{b2m} are yields for beans in the mixture for crops one and two and Y_{b1s} and Y_{b2s} are the corresponding yields for the bean sole crops. For economic values, one may use the ratio or difference of $[p_c Y_{cm} + p_b(Y_{b1m} + Y_{b2m})]$ and $[p_c Y_{cs} + p_b(Y_{b1s} + Y_{b2s})]$ to compare mixture yields with sole crop yields. p_c and p_b represent the value or profit per unit of yield for cassava and beans, respectively. Melon may be used as a cover crop to control erosion during the first part of the cassava growing season.

19.3 Ecological Studies

Many of the procedures described in Chapters 15–18 are directly useful for ecological investigations. These studies involve plant associations and competition. General mixing ability and specfic mixing ability concepts are useful in studying plant associations. The ideas related to competition and measures of competition are intimately related to ecological investigations. Plant communities are composed of many and varied species, and their population numbers are influenced by the type and number of plants and animals present. The makeup of a plant community is greatly influenced by the temperature, the amount of moisture, topology types, soil types, and other weather and topological related characteristics. Many and varied relationships exist in the establishment and maintenance of a plant community. Ecological sytems are dynamic and evolving over time.

19.4 Golf Courses

Golf course establishment and maintenance involves plant populations for the tees, for the fairways, for the rough, and for the greens. This will continue to be true until, if ever, artificial turf golf courses are constructed. The type of grass and legume mixtures varies for each of the four components. The particular mixtures selected for a course will depend on many things such as cost, moisture, temperature, soil type, fertilizer usage, and traffic on the various parts of the course. In addition, each of the 18 or so holes may require something different in order to meet the standards set for the course. Par-three hole tees will have many more divots than will par-four and par-five holes and thus may require different grass mixtures. A response that is desired is "condition" of tee, fairway, rough, and green for each hole as determined by the course manager. Mixtures requiring minimum care will be less expensive to maintain and, hence, will make the course more profitable.

19.5 Medicine and Pharmacology

An LER is a measure of land use. This measure may also be used as a measure of material use in the pharmaceutical production of drugs. One method for doing this is to include sequences using drugs A, B, and C, say, on a continuous basis (continuous cropping in agricultural terms). Rather than using the term land equivalent ratio, we denote this as a *material equivalent ratio* (MER). For example, suppose drugs A, B, C, and D are useful in treating a particular condition, and it is desired to know if mixtures of two, three, and four drugs will enhance drug effectiveness over using the single drugs (sole crops). If it does, then the quantity of drugs in a dosage can be reduced, thereby saving or making more efficient use of the material available. An MER greater than unity is required if there is to be a savings in materials. If a 50–50 mixture of two drugs is more effective than either drug alone, then the MER is greater than 1 and less material for the mixture will be required to attain the same potency.

The concepts of general mixing ability, bi-specific mixing ability, tri-specific mixing ability, etc. are pertinent to clinical trials on animals or people. For example, drugs A, B, and C and the two-drug mixtures AB, AC, and BC may all be noncarcinogenic, but the three-drug mixture ABC could be carcinogenic. Here, the tri-specific mixing effect is important. In order to protect the participants in a clinical trial, a participant may be given the individual drugs A, B, and C in periods 1, 2, and 3 in random order. Then, if there are no adverse reactions, the two-drug mixtures AB, AC, and BC are given in periods 4, 5, and 6 in random order. The three-drug mixture is not given until period 7. In experiments of this type, it is necessary to determine that this method of presenting the treatments to the participants does not affect the results by introducing biases. Some participants with low drug tolerance levels may not be able to complete the seven-period sequence of treatments. Previous trials on animals may not have these requirements and this may allow the experimenter to predict whether or not adverse reactions may occur in the trial on humans.

The statistical designs and analyses presented in Chapters 15 and 16 are directly usable in pharmaceutical studies involving mixtures of drugs, vitamins, supplements, etc. Almost all individuals using these items use more than one drug and thus are using a mixture. Some individuals are using 10–20 different drugs and vitamins. Their adverse effects on the human body are usually unknown if the effects are minor. Only major adverse and short-term effects are readily detectable. Long-term effects on the kidneys, for example, may go undetected. Here, again, experiments on animals may throw light on whether or not adverse effects exist. In animal experiments, there may be no sequencing of treatments as described in the preceding paragraph. Instead of using a repeated measures design, the individual drugs and mixtures may be applied in random order and perhaps to different sets of animals, but long-term effects in this type of experiment go undetected. Even though this design is used, it is recommended that the design to be used on humans be tried on animals first. There can be cooperative, inhibitive, or neutral effects of drugs on each other in mixtures. *Assuming no interaction* among drugs can prove disasterous to the user.

A recent drug treatment involves giving "drug cocktails," which are mixtures of drugs. One such example is the drug cocktail used in the treatment of HIV-infected individuals. The mixture of drugs slows the progress of the disease much better than any of the components of the cocktail, indicating positive combining ability (interaction) of some or all of the drugs. The drugs in the mixture may enter in the amount used for individual (sole) drugs, they may enter the cocktail in a replacement series style, or they may enter in some other amounts.

The different types of competitive effects described in Chapter 17 are discussed in medical literature but under different names. The terminology used here is neutral, inhibitive, and cooperative effects and the ideas used in medical literature are identical to those used herein. This is another example of using the same idea or procedure in two different subject matter areas and using entirely different names for them.

Example 19.1. Federer (1991) presents a discussion of an experiment conducted by Jellinek (1946) which involved a replacement series design of the following nature:

Treatment A—the commercial drug as formulated with compounds $a + b + c$
Treatment B—compound a and c with b omitted
Treatment C—compounds a and b with c omitted
Treatment D—a pharmacologically inactive placebo with all three compounds omitted

Drug A was used in the treatment of headaches. Compounds b and c were in short supply and it was desired to determine if a two-compound mixture was as effective as a three-compound mixture. If B and C were as effective in treating headaches as A, a smaller quantity of compounds b and c would be required. Treatment D was included to distinguish between psychological and physiological headaches. Many drugs may have several active compounds and it may be desirable to determine the effect of eliminating some compounds as done in the above experiment.

Example 19.2. Federer *et al.* (1977) discuss an experiment involving the treatment of patients with induced asthama attacks. Two drugs, A and B, alone and a mixture of the two drugs, $A + B$, were used. A placebo P was included to make four treatments in the experiment. The mixture $A + B$ was not given to a patient until after A and B drug reactions on a patient were noted.

19.6 Education

In obtaining an education, a student has several types of instruction in a variety of subjects, all the way from kindergarten to postgraduate work at a university. The size of the mixture of courses and subjects is often in the four to six range. The sequence of courses as well as the mix has an effect on the progress of the student in digesting the material presented. The type of presentation such as self-paced, intructor paced, auditory, visual, computer aided, and correspondence (distance learning) is another form of mixtures found in education. One or more of the presentation types may be used in each course. Thus, there are two types of mixtures present. The type of presentation mixtures represent a replacement series as discussed in Chapter 17. The subject matter courses have a variety of contents, combinations, and sequences. The sequence of courses may have a profound effect on the rate at which a student absorbs the material. For example, where should a gym class, an aerobics class, or a dance class be placed in order to motivate the student to perform better in the following class? Do students learn faster in an English (or Mathematics) class that meets prior to or after a physically active class? Do students learn faster if an English class is taken as a morning or as an afternoon class? There is also a mixture of topics within a course, e.g., Statistics. Should the first course in Statistics contain only point and interval estimation, with hypothesis testing being relegated to the end of a second course? Should all first courses contain the elements of statistical design, including survey, model, treatment, and experiment design? If hypothesis testing were not in the first course in statistical methods, would it be possible to complete statistical methods up through multivariate analyses (other than multiple regression) in one term?

From the preceding, it is obvious that educational systems are made up of mixtures. Owing to the nature of educational instruction, the class of repeated measures designs (see, e.g., Federer and Kershner, 1998) is important for research on educational systems. With minimal planning, many of the questions could be resolved by conducting the research with ongoing classes. For example, in a high school, there may be multiple classes of four, say, major subjects. A treatment design could consist of different sequences of the four courses. By conducting the experiment over groups of students in the same year or different years, data would be available to determine the sequence of courses resulting in the highest level of achievement. It may also be possible that students could classified into groups which would perform best with different sequences of courses. For three courses, there would be 6 sequences and for 4 courses, say E = English, H = History, M = Mathematics, and S = Social Studies, there would be 24 sequences as follows:

p	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	E	E	E	E	E	E	H	H	H	H	H	H	M	M	M	M	M	M	S	S	S	S	S	S
2	H	H	M	M	S	S	E	E	M	M	S	S	E	E	H	H	S	S	E	E	H	H	M	M
3	M	S	H	S	H	M	M	S	E	S	E	M	H	S	E	S	E	H	H	M	E	M	E	H
4	S	M	S	H	M	H	S	M	S	E	M	E	S	H	S	E	H	E	M	H	M	E	H	E

This is a repeated measures design with $s = 24$ sequences of $v = 4$ treatments (subjects) and $p = 4$ periods. It is a 4-row (period) by 24-column (sequence) design. In such a design, the direct effect and residual effects are estimable. The sequence totals as well as the four treatment means for each period are also of interest.

Another area where questions arise is in the teaching of a second language. Would the student become more proficient in a language if all courses in a given year were taught in that language rather than taking a language course over several years? Would teaching all courses in a year in the second language be deterimental to the learning of the material in the non-second-language courses? What mixture is best educationally and for what objective? Short courses need to have the appropriate mixture of topics in order to be successful. Training courses will need the right mixtures in order to be optimally effective in accomplishing a desired goal. In writing textbooks, the order and makeup of topics is important and contributes significantly to the usefulness for reader comprehension.

As indicated above, replacement series and sequential presentation of courses are important concepts in educational research, just as they are in agriculture. The concepts of general mixing ability and specific mixing ability of combinations and sequences of a combination of courses are integral to educational research. The ideas and procedures of Chapter 15–18 may be applied in this area. The ideas of cooperative, neutral, and inhibitive competition effects of various components of the mixture are useful in interpreting the results of research projects.

19.7 Exercise and Aerobics

Exercise and aerobics programs are a mixture of several procedures. Each instructor selects a mixture thought best for the class. In addition to these mixtures, mixtures of foods and/or supplements are included when exercise is a part of a dieting program. Stretching and other exercises are used before sports practices or events and before dance classes. The particular mixture used is determined for each activity. Frequency and duration of activities need to be considered. An appropriate mixture of calisthenics and pre-practice exercises and practice times for such sports as football may decrease the number of injuries players suffer. In order to keep interest at a high level, aerobics and gym classes need a diverse mixture of routines. Special exercise routines are developed for arthritic, heart, and other patients. A mixture of exercise and nutrition is well known to be beneficial for good health and well-being of both mind and body. Methods of measuring effects of an aerobic or exercise program are often subjective and inferential rather than being based on objective procedures. Sometimes, the effects are small and cumulative, and need to be observed over long periods of time. For example, does a program increase expected life span? How much does a program decrease (increase) cost and frequency of medical treatment in a 25-year period? How long should the treatment period be in order to compare the effects of two competing programs?

19.8 Engineering

Engineering and industrial activities of many and diverse types involve mixtures of products or procedures. In building construction. e.g., the proportion and type of metal reinforcement, sand, water, ash, and cement required to achieve a particular outcome needs to be determined for each construction or building project. Floors, walls, and supports require different mixtures. Bridges, buildings, and highways may have different requirements for a mixture. In the manufacture of automobiles, camshafts, axles, engine blocks, spark plugs, etc. require various alloys made up of mixtures of materials. Too much carbon with iron may make the product too brittle and easy to break. Correct proportions of materials is essential to obtain the desired quality of a product. The entire product, such as an automobile, is a mixture of many products involving many alloys, plastics, cloth, leather, etc. and this mixture must be ascertained in order to obtain the desired product. The roadbeds of highways usually consist of a gravel base overlaid with concrete formed from a mixture of cement, sand, ash, and water. A replacement series experiment may be used to determine the proportion of each of the above four components in the mixture required to obtain concrete of the desired characteristics such as maximum longevity, minimum porosity to moisture, ability to hold up under various types of vehicular traffic, etc. Recently, it was reported that New York highway engineers rediscovered what the Romans knew centuries ago when they built the Colosseum. If ash (about 24%) were added to the cement, sand, and water mixture, the hardness was increased, the porosity to water was reduced, and the longevity of the product was greatly increased. Macadam is a pavement of layers of small stones, gravel, or sand bound together with tar or asphalt. It was discovered by John L. McAdam (1756–1838), a Scottish engineer. This mixture in various forms is a popular form of pavement for highways worldwide. In some areas, the mixture contains ground-up discarded automobile tires. There are experiments being conducted on many types of mixtures for highway construction. The following examples indicate some types of mixtures.

Example 19.3. Box *et al.* (1978) present the following illustration as an example of a replacement series experiment: "To reduce the amount of a certain pollutant, a waste stream of a small plastic molding factory has to be treated before it is discharged. State laws require that the daily average of this pollutant cannot exceed 10 pounds. The following 11 experiments were performed to determine the best way to treat this waste stream":

Order of experiment	Chemical brand	Temperature (deg. F)	Stirring	Pollutant (lb./day)
5	A	72	None	5
6	B	72	None	30
1	A	100	None	6
9	B	100	None	33
11	A	72	Fast	4
4	B	72	Fast	3
2	A	100	Fast	5
7	B	100	Fast	4
3	AB	86	Intermediate	7
8	AB	86	Intermediate	4
10	AB	86	Intermediate	3

The AB was a 50–50 combination of chemical brands A and B as would be used in a replacement series treatment design. The 86°F temperature is intermediate between 72 and 100. It would have been informative if the experimenter had added the four combinations of treatment AB at (72°F, none), (72°F, fast), (100°F, none), and (100°F, fast). It is possible that AB at 72°F with no stirring would have been satisfactory. This would save energy in temperature and in stirring, and if brand B were cheaper, this combination would be selected. If AB was effective as a mixture, a replacement series experiment needs to be conducted to determine the optimum proportion of the two ingredients.

Example 19.4. Anderson and McLean (1974) present the following example measuring stress–rupture life of four alloys (mixtures) used to make turbine blades in fanjet engines. Three blade temperatures during take-off of aircraft were used:

Alloy	Temperature 1	Temperature 2	Temperature 3	Total	Mean
a	185	182	182	549	183.0
b	175	183	184	542	180.7
c	171	184	189	544	181.3
d	165	191	189	543	181.0
Total	696	740	744	—	—
Mean	174	185	186	—	181.7

The four alloys gave essentially the same means when averaged over the three temperatures as did temperatures 2 and 3 when averaged over alloys. One linear contrast accounting for a good share of the variation is the contrast of temperature 1 with the average of temperatures 2 and 3. A second linear contrast which accounts for a large share of the remaining variation is the contrast of the linear regression coefficient of stress–rupture on alloy for temperature 1 with that from temperatures 2 and 3. These two single degree of freedom contrasts account for most of the variation in this example. The reader may wish to verify this by performing the necessary computations.

Example 19.5. Anderson and McLean (1974) discuss a situation wherein the time in seconds of disintegration of 4 magnesium trisilicates for each of 10 tablets is obtained. The four mixtures are

A = magnesium stearate (material going through 16 mesh)
B = talc powder (material going through 16 mesh)
C = liquid petrolatum (material going through 16 mesh)
D = magnesium stearate (material going through 20 mesh)

The results were

	A		B		C		D	
	20	42	8	12	50	124	151	178
	28	25	10	24	67	72	125	151
	36	24	12	10	90	78	180	152
	16	31	16	19	103	70	140	161
	25	33	9	10	90	76	175	118
Mean		28		13		82		154
Variance		60		26		431		436

Note the inequality of variances. Each of A, B, C, and D is a mixture of components which needs to be determined in order to have a product that disintegrates in the required time period.

Example 19.6. As explained in Chapter 17, linear programming involves determination of a mixture to optimize cost, material, labor, etc. To illustrate, suppose that two refineries make different amounts of three grades of gasoline, A, B, and C, during a single run so that their amounts are in a fixed proportion. Since the particular proportions made do not correspond with the needs of a consumer, it has to be decided how many runs to obtain from each refinery in order to meet the requirements. The data might be of the following nature:

Gasoline grade	Refinery 1	Refinery 2	Needs of consumer
A	1	1	100
B	3	4	340
C	1	5	150
Cost/run	\$300	\$500	

How many runs should be obtained from each refinery in order to meet the requirements and minimize cost?

Example 19.7. To control a certain crop disease, it is necessary to use 6 units of chemical A, 10 units of chemical B, and 8 units of chemical C, i.e., in the ratio 6:10:8. Each of two different brands is sold in barrels containing the following proportions:

Chemical	Barrel 1	Barrel 2	Required proportions
A	3	1	6
B	3	3	10
C	3	4	8

What mixture of barrels 1 and 2 should be used to obtain a 6:10:8 = 3:5:4 mixture of the three chemicals?

Example 19.8. Two factories produce low, medium, and high grades of a product in a daily run. An order specifying the number of tons of the three grades is received. The question is to determine how many days to run each factory to fill the order. The number of tons of each grade produced each day and the cost of a day's run are as follows:

Grade of product	Factory 1	Factory 2	Tons ordered
Low	8	2	16
Medium	1	1	5
High	2	7	20
Cost/day	\$1000	\$2000	

The most economical plan is to run factory 1 for 3 days and factory 2 for 2 days. This mixture of runs for the two factories produces the necessary amount with a 12-ton surplus of the low grade, but it is the most economical when using an entire day's output.

Example 19.9. An office manager has two groups performing three different tasks in differing amounts. An order requiring different numbers of the three tasks is received. The manager needs to decide to put group 1 on the job for x_1 days and group 2 on the job for x_2 days. The situation is as follows:

Task	Group 1	Group 2	Number of tasks
A	8	2	16
B	1	1	5
C	2	7	20
Minutes required	1000	2000	

Since the numbers are the same as the previous example, $x_1 = 3$ days for group 1 and $x_2 = 2$ days for group 2 will minimize the number of full days required to handle the number of tasks.

Example 19.10. In the manufacture of marker flares, ingredients like magnesium, sodium nitrate, strontium nitrate, and binder material are mixed together in varying proportions to obtain the desired characteristic of the marker flare. A replacement series experiment is conducted to obtain the desired product for visibility, for brightness, for length of burning time, etc.

Example 19.11. In the production of stainless steel, a mixture of iron, carbon, chromium, and other material is used. A replacement series experiment is conducted to determine the desired property of the product such as resistance to staining, flexibility, brightness, and other characteristics.

19.9 Marketing and Transporation

Grocery, shoe, clothing, hardware, drug, electrical, houseware, lawn, furniture, etc. stores sell a variety of products and often several brands of the same product. Shelf and floor space are usually at a premium, and the store manager is faced with the problem of how much shelf or floor space to devote to a product. The problem of where to locate the product as well as which other products to display with it must be determined for each mixture of products. In order to effectively promote one product, it is necessary to determine where and how to display the product. The appropriate mixture of adjacent products is required in order to maximize sales. Experiments in grocery or drug stores will usually be using four to six shelves for an experimental display. Optimal experiment designs to measure horizontal, vertical, and diagonal competing effects of other brands on a product are yet to be developed. Professor D. Raghavarao, Temple University, and the author have constructed a particular design for this situation, but no general class has been constructed.

In the transportation of automobiles from the manufacturer to the dealer, automobile-carrying trucks are built to carry a certain mixture of types of cars and trucks. To minimize transportation costs and to meet the orders from dealers in an area, a mixture of types of trucks will often be required. Moving vans face similar situations in that they are often carrying a variety of goods and have several destination points. All types of transportation have a mixture of activities and goods, and in order to be profitable and stay in business, costs need to be minimized.

Example 19.12. Raghavarao and Wiley (1986) discuss an experiment design for an experiment on soft drink preferences when the customer has a choice of one of four soft drinks. Combinations of four of eight soft drinks were used in the experiment. The eight soft drinks were

A = regular Coke	B = diet Coke	C = regular Pepsi	D = diet Pepsi
E = regular 7-Up	F = diet 7-Up	G = regular Sprite	H = diet Sprite

The scores and basic 3-design with parameters $v = 8$, $b = 14$, $k = 4$, $r = 7$, $\lambda = 3$, and $\delta = 2$ are as follows:

Set	A	B	C	D	E	F	G	H
1	10	1	—	7	—	—	—	6
2	—	2	10	—	11	—	—	1
3	—	—	16	4	—	1	—	3
4	—	—	—	4	9	—	6	5
5	8	—	—	—	6	4	—	6
6	—	5	—	—	—	7	10	2
7	6	—	9	—	—	—	3	6
8	—	—	10	—	4	5	5	—
9	8	—	—	3	—	5	8	—
10	10	6	—	—	6	—	2	—
11	7	6	8	—	—	3	—	—
12	—	6	5	7	—	—	6	—
13	7	—	6	3	8	—	—	—
14	—	9	—	7	5	3	—	—
Totals	56	35	64	35	49	28	40	29

They present various types of analyses and demonstrate what happens when some of the soft drinks are removed from the shelves.

Example 19.13. A television station has two types of programs with each type attracting different numbers of viewers. An advertiser puts certain requirements on the proportion of time given to music and to advertisements as follows:

Item	Program type A	Program type B	Requirement
Music	20 min	10 min	80 min
Commercial	1 min	1 min	6 min
Number of viewers	30,000	10,000	

How many times should program type A and how many times should program B be run in order to meet the requirements and maximize the number of viewers? A and B are mixtures as are the mixtures of the two program types.

Example 19.14. A trucker has a mixture of 900 boxes of oranges, 700 boxes of grapefruit, and 400 boxes of tangerines. Each possible market city has a different mixture of prices for the fruits. Given the following prices per box, a city is selected as the destination:

City	Orange	Grapefruit	Tangerine
New York	4	2	3
Cleveland	5	1	2
St. Louis	4	3	2
Oklahoma City	3	2	5

The answer is St. Louis when the distance to market is the same.

Example 19.15. A trucking organization has three trucks which can carry three types of machines but in different numbers. The minimum number of trips for each truck in order to deliver a request for machines of the three types needs to be determined. In tabular form, the problem is as follows:

Machine	Truck 1	Truck 2	Truck 3	Request
A	1	1	1	12
B	0	1	3	10
C	2	1	1	16
Machines/load	3	3	4	

19.10 Nutrition and Diet

A variety of foods, breakfast cereals, breads, meats, vegetables, etc., are used in diets with very few if any diets consisting of a single food. Mixtures of foods in diets vary with activity, age, sex, health, race, and other attributes. The content of a food is made up of many components such as fats (saturated and unsaturated), minerals, carbohydrates, vitamins, and protein. An appropriate mixture of these components is needed to determine the amount of food recommended for consumption in a diet. In many cases, the amount of many of the components are listed on the package cover. Different brands of food, say breakfast cereal, vary widely in amount of the various components. A consumer may decide that in order to obtain the desired amounts of each of the components, a mixture of breakfast cereals, say, is required. Diets for certain medical situations may need to be determined, as, for example, in the coagulation time for blood, for diabetics, for intestinal disorders, and for other maladies.

Example 19.16. Four common breakfast cereals and their listed ingredients are

Frosted Shredded Wheat	Raisin Bran	Rice Krispies	Corn Flakes
Whole wheat	Whole grain wheat	Rice	Milled corn
Sugar	Raisins	Sugar	Sugar
Brown sugar	Wheat bran	Salt	Salt
Gelatin	Sugar	Corn syrup	Malt flavor
	Salt	Malt flavor	Corn syrup
	Wheat flour		
	Malted barley		
	Honey		
	Corn syrup		

Four to nine ingredients enter into the mixture called a breakfast cereal. Some nutritional facts listed for these four breakfast cereals are as follows:

Cereal	Fat	Protein	Carbohydrate	Fiber	Calories	Sodium
Frosted Shredded Wheat	2%	4 g	15%	20%	190	0%
Raisin Bran	2%	4 g	16%	31%	190	13%
Rice Krispies	0%	2 g	10%	0%	160	15%
Corn Flakes	0%	2 g	8%	4%	100	13%

The percentages are of daily requirements from one serving (cup) of cereal. Protein is given by number of grams. For calories, the first two have almost twice as much as Corn Flakes. One of the boxes for the cereals advertised that 13 “essential” minerals and vitamins, and another said 10, were obtained from each serving. Many combinations of ingredients needed to be tried before arriving at the present mixture. Designs and analyses like those in previous chapters are useful in arriving at the final product. Likewise, many combinations needed to be investigated to obtain the desired nutritional amounts, e.g., 0% sodium or 0% fat.

19.11 Comments

As illustrated for many areas, the use of mixtures is a widespread phenomenon. The determination of which mixture to use involves many and diverse methods and many and diverse characteristics. The methods and procedures discussed in this book can be used profitably in many areas. New procedures and methods may, and probably will, need to be developed for many situations involving the use and choice of mixrures. A mixture is a system and users must consider what happens to the entire system and not just to each component individually. Consideration of a part of the system and ignoring other parts of the system can lead to failure of the system. The interactions of the components of the system must be considered, as some are beneficial and others are detrimental to the system.

A factorial treatment design is one form of mixture experiment which involves n factors, with each factor having k_i levels for factor i. This type of mixture has been well discussed in the literature. The type of mixtures discussed in this book involves n factors in the mixture with one prescribed level. A replacement series is a particular kind of mixture wherein the addition of an amount of a factor means other factor amounts must be decreased. Some of the intercropping procedures involve a replacement series, but others do not. The factorial ideas do not apply to the type of mixtures discussed herein. Replacement series ideas as discussed by Cornell (1990), e.g., do not cover all the situations encountered in intercropping experiments. The many kinds of mixtures found in the universe of mixture systems may even need ideas beyond any of those discussed herein. However, if any procedure for any type of mixture is usable, it should be used. From the bibliography discussed in Chapter 20, it is obvious that several of the procedures discussed in this book could be used beneficially to summarize results from experiments.

19.12 Problem

19.1 List other examples and other areas where mixtures and/or systems of treatments are used.

19.13 Literature Cited

Anderson, V. L. and R. A. McLean (1974). *Design of Experiments — A Realistic Approach*. Marcel Dekker, Inc., New York.

Box, G. E. P., W. G. Hunter, and J. S. Hunter (1978). *Statistics for Experimenters — An Introduction to Design, Data Analysis, and Model Building*. John Wiley & Sons, New York, pp. 443-444.

Cochran, W. G. (1939). Long term agricultural experiments. *J. Roy. Statist. Soc., Series B* **6**, 104-148.

Cochran, W. G., K. M. Autrey, and C. Y. Cannon (1941). A double change over design for dairy cow feeding experiments. *J. Dairy Sci.* **24**, 937-951.

Cornell, J. (1990). *Experiments With Mixtures: Designs, Models, and Analysis of Mixture Data*. John Wiley & Sons, Inc., New York.

Crowther, F. and W. G. Cochran (1942). Rotation experiments with cotton in the Sudan Gezira. *J. Agric. Sci.* **32**, 390-405.

Federer, W. T. (1991). *Statistics and Society — Data Collection and Interpretation*, 2 edn., revised and expanded. Marcel Dekker, Inc., New York, Chapter 8.

Federer, W. T. and G. F. Atkinson (1964). Tied-double-change-over designs. *Biometrics* **20**, 168-181.

Federer, W. T. and R. Kershner (1998). Comparison of classes of changeover designs. *J. Combin. Inform. Syst. Sci.* **23**.

Federer, W. T., F-C. H. Lee, and W. R. Holt (1977). Statistical analyses for a set of medical data. BU-628-M, Technical Report Series of the Biometrics Unit, Cornell University, Ithaca, NY.

Kass, D. C. L. (1978). *Polyculture cropping systems: Review and analysis*. N.Y. State University Agriculture Life Science, Cornell Intl. Agric. Bull. 32.

Jellinek, E. M. (1946). Clinical tests on comparative effectiveness of analgesic drugs. *Biometrics* **2**, 87-91.

Patterson, H. D. (1953). The analysis of the results of a rotation experiment on the use of straw and fertilizers. *J. Agric. Sci.* **43**, 77-88.

Patterson, H. D. (1959). The analysis of a non-replicated experiment involving a single four course rotation of crops. *Biometrics* **15**, 30-59.

Patterson, H. D. (1964). Theory of cyclic rotation experiments (with discussion). *J. Roy. Statist. Soc., Series B* **26**, 1-45.

Raghavarao, D. and J. B. Wiley (1986). Testing competing effects among soft drinks. In *Statistical Design: Theory and Practice* (Editors: C. E. McCulloch, S. J. Schwager, G. Casella, and S. R. Searle), Proc., Conference in Honor of Walter T. Federer, Biometrics Unit, Cornell University, Ithaca, NY, pp. 161-176.

Williams, E. J. (1949). Experimental designs balanced for the estimation of residual effects of treatments. *Australian J. Sci. Res.* **2**(A), 149-168.

Yates, F. (1949). Design of rotation experiments. Commonwealth Bureau Soil Sci., Tech. Comm. No. 46.

CHAPTER 20

An Intercropping Bibliography

Literature citations on intercropping studies date back many centuries. Early references are given in The Holy Bible, Leviticus 19:19 and Deuteronomy 22:9–11. Since intercropping in one form or another dates back to antiquity, there are no doubt many references to intercropping over the ages. A bibliography for twentieth century literature was prepared. More than 3000 literature citations appear in the bibliography. Owing to the large amount of text space this would require, it was decided not to include it in the present volume. Instead, copies of the bibliography are available as follows:

(i) Hard copy of 124 pages.
(ii) Disk copy in Microsoft Word.
(iii) Biometrics Unit, Cornell University web site.

Copies of (i) and (ii) are available from the Biometrics Unit, 434 Warren Hall, Cornell University, Ithaca, New York 14853, at a cost of $7.50 for (i) and $5.00 for (ii). These costs are to recover the cost of reproduction and mailing. The web site for the Biometrics Unit is http://biom.cornell.edu and the e-mail address is biometrics@cornell.edu.

As may be noted from the bibliography, there has been considerable research activity on the many and diverse aspects of intercropping (see Chapter 19). It would appear that many of these studies could have profitably used some of the statistical design and analyses given in this volume in order to extract additional information from their experimental data. This would have allowed for a more efficient use of research resources. For example, the procedures given in Chapter 18 would have been useful in extracting information for many of the studies.

The procedure followed to obtain the bibliography was as follows:

(i) A computer search of the literature was made in 1983.
(ii) A computer search of the literature was made in July 1997 using the Cornell University Gateway Library Resource Agricola.
(iii) A number of doctoral dissertations were obtained and literature citations obtained from them.
(iv) Many reprints, technical reports, and papers on intercropping were obtained and these were used to supplement references obtained from other sources.

The overlap of references in (i) and (ii) was minimal. The references from (ii) represented about one-half of the total references, indicating that computer searches to date will not fully recover all citations. Also, despite the extensive search made here, there are always some references that were missed, as is the case with any bibliography. In compiling the bibliography, it was noted that there were errors in citations given by the various authors. Efforts were made to obtain correct citations, but it is possible that errors still exist. In compiling bibliographies, it has been noted that authors of scientific articles frequently make mistakes in their literature citations, probably because this is considered the least important part of their papers. Some references were omitted when the citation was incomplete and could not be verified.

Index