ON THE DESIGN AND ANALYSIS OF INTERCROPPED EXPERIMENTS

by

Walter T. Federer and Anila Wijesinha

BU-735-M *

May 1981

1. Introduction

Intercropping, the growing of two or more crops simultaneously on the same area of land, is an age-old agricultural procedure, especially in the tropics. Experiments on intercropping involve many statistical design and analysis difficulties not encountered in sole or single crop experiments. The problems of fixed-ratio mixtures have been largely neglected in Statistics except in the areas of diallel crossing systems and paired comparisons; these are special cases of fixed-ratio mixture designs where the ratio of the pair in a mixture is 1:1. It should be noted that there is a large literature on mixture designs for which it is desired to estimate a ratio giving some desirable property; here the ratio is prescribed for the designs we are considering. It should be further noted that fixed-ratio mixture designs are important in ascertaining effects of a chemical compound, of a diet, of a recreational program, of an educational program, of combination of drugs, of combining questions in a survey, etc. Much thought, creativity, and work will be required to solve the statistical design and analysis problems for fixed-ratio mixture experiments.

In the next two sections we consider some treatment and experiment design problems for intercropping experiments. Then, we consider statistical analyses for four particular situations involving mixtures of two crops. Some involve modeling considerations. Then, we consider more than two crops in a mixture. Because of the increasing difficulties associated with three or more crops in a mixture, it is highly desirable to first be well acquainted with designs and analyses for mixtures of two crops. A number of different univariate and multivariate statistical analyses are considered. The advantages and disadvantages of each are discussed with special attention being given to difficulties with

---

* Paper in the Mimeo Series of the Biometrics Unit, Cornell University, Ithaca, New York.

multivariate analyses procedures. Finally, a set of selected references is given.

2. Treatment Design

One goal of most intercropping experiments is to assess the yield, or other response, of the mixture relative to what the crops in the mixture would have done when grown as sole crops. In order to do this, sole crop treatments, as well as the mixture, must be included in the selection of treatments, the treatment design. When one is interested in general mixing and specific mixing effects (see Federer 1979), various types of designs are possible for v cultivars in mixtures of size k in equal proportions; the treatment design depends upon how many and what type of effects are to be estimated. For example, if v = 7, k = 3, and only general mixing effects are to be estimated, then the following treatment design suffices (numbers refer to cultivars 1 to 7):

mixtures (design I)

1	2	3	4	5	6	7
2	3	4	5	6	7	1
4	5	6	7	1	2	3

or, alternatively:

mixtures (design II)

3	4	5	6	7	1	2
5	6	7	1	2	3	4
6	7	1	2	3	4	5

If one wishes to estimate general mixing and specific mixing effects for two cultivars (bi-specific mixing effects), 21 mixtures as follows are required:

mixtures (design III)

1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	3	3	3	3	4	5
2	2	2	3	3	4	4	4	6	3	3	4	4	5	5	4	4	5	6	5	6
3	5	7	5	6	5	6	7	7	4	6	6	7	6	7	5	7	7	7	6	7

If one wishes to estimate general mixing effects free of cultivar effects as sole crops, it is necessary to include the seven sole crops in the treatment design. If one wishes to estimate tri-specific mixing effects, the interaction effect of a specific mixture of three cultivars, in addition to the other above effects, it is necessary to use Designs I, II, and III as well as the sole crops. Thus, the treatment design would consist of all possible mixtures (35) of size three plus the seven sole crops, or v = 42 treatments.

Each set of goals requires a specific treatment design. It is imperative to coordinate the goals of an intercropping experiment with the treatment design if one is to have a successful experiment.

3. Experiment Design

The experiment design (arrangement of treatments in an experiment) for the v treatments of the previous section may be a completely randomized, randomized complete block, incomplete block, split-plot, split-block, row by column, etc. design. The experiment design is selected to control heterogeneity in the experimental area and for the treatment contrasts of most importance. After blocking in such a manner as to obtain relatively homogeneous experimental units within blocks, the experimenter decides which treatments are of most importance and whether to use complete confounding (split-plot and split-block) or partial confounding of particular treatment contrasts (incomplete block designs). When v is relatively large, some type of incomplete block design may be necessary. Certain types of statistical analyses may require specific experiment designs.

4. Statistical Analyses for Intercropped Experiments

There are many types of intercropped experiments and many different goals. Hence, it is profitable to delineate some of the types and to give analyses for each type and goal. It is important to understand the concepts and analyses on pairs of intercrops before proceeding to triplets, quartets, etc. Also, we have considered the following types of intercropped experiments:

(i) one main crop intercropped with supplemental crops,

(ii) two or more main crops form the intercrop mixture,

(iii) (i) and (ii) with density per hectare as a variable,

(iv) modeling considerations when intercrop yields are separable,

(v) modeling considerations when intercrop yields are not separable, and

(vi) variable density and spatial arrangements.

4.1. One main crop with supplemental crops

Statistical analyses for this situation differ little from standard statistical procedures. The goal is to obtain responses for the main crop which are not less than ten percent, say, of the response for sole crop yields. The supplemental crops are just supplemental to the yields of the main crop. A farmer might intercrop corn and cucurbits, but his interest would be in the yield of corn. Statistical analyses would be for the yield of the main crop.

4.2. Two or more main crops form the intercrop mixture

This situation makes the statistical analyses exceedingly difficult over the previous situation. We must now consider whether we want:

(i) statistical analyses for each main crop, or

(ii) statistical analyses of responses for all crops in the mixture.

The former situation does not present too many difficulties when density remains constant and when the responses are for comparative purposes only. Standard statistical analyses are sufficient to handle many situations. It should be noted that situation (i) is the one considered in the majority of papers published on intercropped experiments.

When one considers situation (ii) the immediate question is which joint response of the yields of two or more crops is to be considered. Some contenders that have come to mind are:

(i) total yield for the crops in the mixture,

(ii) total calories for the crops in the mixture,

(iii) total protein for the crops in the mixture,

(iv) total value of the crops in the mixture,

(v) total profit of the crops in the mixture,

(vi) total value of the crops in the mixture to the farmer,

(vii) fertilizer replacement value,

(viii) insect and disease control value,

(ix) a multivariate index,

(x) the availability of produce for food and market throughout the year,

(xi) land equivalent ratio,

(xii) stability of yields,

(xiii) etc.

Most of the above involve univariate statistical analyses. Standard statistical procedures are sufficient in most cases once one has a joint response variable to analyze. We have found it expedient to obtain statistical analyses for several of the above joint responses, and to use these for comparative purposes to determine mixtures resulting in highest values for the joint response variable. There are inherent statistical difficulties with many of the above variables. Several may involve heterogeneity of error variances, while others like land equivalent ratios (see, e.g., Mead and Stern (1980) and Mean and Willey (1980)) and stability measures result in dependent observations

The naive statistician usually says this appears to be a multivariate analysis problem, and the entire problem should be treated as such. Multivariate analyses procedures can be used to a limited extent, but they must be used with caution. The individual crops responses are treated as the multivariate variables in Pearce and Gilliver (1978, 1979) and Wijesinha, et al. (1980). However, as Mead and Riley (1981) point out, there are some major disadvantages with the method for mixtures of two crops; for example:

(i) comparisons of crop yields in a mixture with sole crop yields present many difficulties and impossibilities,

(ii) estimation is required rather than graphical plots and significance testing, and

(iii) the fundamental assumption of constant correlation between the yields for the two crops is often violated, especially with varying densities and spatial arrangements.

When one considers mixtures of size k of v cultivars, there is the question of whether one can perform a multivariate analysis. Investigations are being made into situations wherein a multivariate analysis can be performed. Another disadvantage of present multivariate analysis theory is that it does not take the experiment design into account and does not relate certain univariate analyses to multivariate analyses. To illustrate, consider a situation wherein m mulching treatments are in a randomized complete block design. Consider c crops as split plots to mulching treatments to obtain a split-plot design, <u>or</u> consider the c crops as stripped to the mulching treatments to obtain a split-block design. If one uses the crop responses as the multivariate responses, an idential multivariate analysis of variance (MANOVA) is obtained for both designs. Since the univariate analyses of variance for the two experiment designs are quite different, how does this show up, or does it, in the MANOVA? What are the relationships for multivariate and univariate analyses in a variety of experiment designs? How robust are multivariate analysis procedures to outliers, nonnormality, variable correlations, variance heterogeneity, etc.? All of these are encountered in experimental investigations, particularly in the tropics. After these procedures have been used extensively and a large body of data built up one will be able to asses the usefulness of multivariate analysis procedures for mixtures of two, three, or more crops in the mixture.

4.3. <u>Modeling considerations when intercrop yields are available</u>

Various response models require consideration in understanding and interpreting the results from intercropping. We have considered fitting the simplest possible models for the responses that would enable intercropping system results to be biologically explained in terms of competition and compensation among the different crops in the mixture. A model whose parameters reflected the complex underlying biological structure of the intercropping system could be very useful in pinpointing types of crops and species of a particular crop, as well as crop combinations that benefitted from being grown as intercrops. Further, since the model would break down the total yield from a crop into relevant components, a better concept of the system could be grasped than if the analysis were only performed on yields using the usual analysis of variance structure.

These abstract ideas may best be illustrated by the use of two examples of possible models, in different intercropping contexts:

a) We consider an intercropping system where each of a number of genotypes of a single crop are to be evaluated on their performance when grown in a mixture with a single other genotype. We assume that for each genotype, the planting density at which yield is optimized in monoculture is known (and not necessarily equal for all genotypes). The performance of each genotype in mixtures is to be compared with its own performance in monoculture, and the performance of the other genotypes both in monocultures and mixtures. We consider a model for yield (or other observed variable) which could be specified in terms of a general mean for all genotypes (since these are genotypes of a single crop, the concept of an overall mean is acceptable), a particular genotype effect over and above the overall mean, a general mixing effect for each genotype, which would be an indicator of its overall ability to mix, and a specific mixing effect due to one particular genotype on another, which would indicate the ability of one particular genotype to mix with another particular genotype.

We consider a system where n genotypes are grown in monoculture at their optimum yielding densities and in all possible pairwise combinations, each at 50% of their monoculture densities. The model response equations for this system could be given by

$$Y_{iim} = \mu + \tau_i + \epsilon_{iim} \qquad \forall\ 1 \le i \le n$$ for monocultures (where suffix m denotes monoculture)

and

$$Y_{i(j)b} = \tfrac{1}{2}(\mu + \tau_i + \delta_i) + \gamma_{i(j)} + \epsilon_{i(j)b}$$ for mixtures

$$Y_{j(i)b} = \tfrac{1}{2}(\mu + \tau_j + \delta_j) + \gamma_{j(i)} + \epsilon_{j(i)b}$$ (where suffix b denotes biblends)

$$\forall\ 1 \le i,\ j \le n \qquad i \ne j\ .$$

Here

Y_{iim} is the yield of the ith genotype in monoculture;

$Y_{i(j)b}$ is the yield of the ith genotype when grown with the jth genotype in a mixture as specified above;

μ is an overall mean effect;

τ_i is the effect due to the ith genotype $\left(\sum_{i=1}^{n} \tau_i = 0\right)$;

δ_i is the general mixing ability of the ith genotype $\left(\sum_{i=1}^{n} \delta_i\right.$ is not necessarily zero);

$\gamma_{i(j)}$ is the specific mixing ability of the ith genotype when grown with the jth genotype $\left(\sum_{\substack{j=1 \\ j \ne i}}^{n} \gamma_{i(j)} = 0 \quad \forall\ 1 \le i \le n\right)$;

and

ϵ_{iim}, $\epsilon_{i(j)b}$, $\epsilon_{j(i)b}$ are random components of variation.

Parameter estimates could be obtained for each of the above with corresponding variances, and an assessment of the τ_i's, the δ_i's and the $\gamma_{i(j)}$'s would give a very clear picture of how each genotype was performing. For example, if $\delta_k - \bar{\delta} \gg 0$ for a particular k and $\tfrac{1}{2}(\delta_k - \bar{\delta}) + \gamma_{k(j)}$ was also positive for all $j \ne k$, this would indicate that the kth genotype did well as a mixer. $\bar{\delta}$ being positive would imply that the crop in general did well as an intercrop with different genotypes, while $\bar{\delta} < 0$ would imply that the crop probably performed better in monoculture.

Variations on this theme could be used to encompass different percentage density combinations of genotypes, but the main point is that the general idea of assessing competitive ability in terms of the δ's and γ's could be a very useful analysis.

b) Another example would suffice to emphasise the usefulness of these types of models. Here we consider intercropping systems involving pairs of different crops, at varying densities. In this case, we assume that a yield density relationship exists and its functional form is known for each crop.

Thus in monocultures, for each crop i $(1 \le i \le n)$ we have for the yield Y_{iim},

$Y_{iim} = f_i(d) + \epsilon_{iim}$ where f_i is the functional form of the ith crop as a function of density d ,

(e.g.: For Y_i increases linearly with density, $f_i(d) = \beta_{0i} + \beta_{1id}$) . We consider the yield of the ith crop when grown with the jth crop at densities d_i, d_j, respectively, as given by

$$Y_{i(j)b} = f_i(d_i) + \gamma_{i(j)}(d_i, d_j) + \epsilon_{i(j)b} \quad ,$$

where $\gamma_{i(j)}(d_i, d_j)$ is the effect due to the interaction of the two crops at those particular densities on crop i . To complete this model, an associated variance structure would be required and then generalized least squares theory would be used to obtain parameter estimates which would describe the performance of the cropping systems.

These are just two ideas of the many possible model structures in this context. We feel that this area should be further explored and applied to the evaluation of intercropping systems. Needless to say, the complexities of obtaining suitable models and associated variance structures for different density and spatial arrangements does not make it easy, and the fundamental concepts underlying any model that is used need to be very carefully considered before any model is applied to the data.

4.4. Modeling considerations when intercrop yields are not separable

In many situations, k out of v stimuli will be applied and a single response will be elicited. If the responses of cultivars, e.g., varieties of white field bean, cannot be obtained separately, then one needs to develop response models for dealing with this situation. For mixtures of two cultivars, one can use many of the concepts, models, and statistical analyses developed for diallel crossing experiments (see, e.g., Eberhart and Gardner (1966), Federer (1979), and Federer *et al.* (1981)). For more than two cultivars in a mixture, it was necessary to extend these concepts and results (Hall (1976) and unpublished papers). The statistical analyses and interpretations become more complicated and complex.

4.5. Varying densities and spatial arrangements

Modeling yields as a function of density is a helpful process for interpreting results from sole crop and intercropped experiments. We have striven for the simplest possible response model equations. For the experimental results on which they have been applied, reasonably good fits to the data have been obtained. Since a true response model equation is often not obtainable, one seeks reasonably good approximations. Our equations are of this nature for densities that cover a relatively short span of the total possible densities. They can be easily adapted to cover a much broader range of densities by introducing another density parameter.

One problem is how to interpret yields from different (or the same) densities in sole crop and in an intercrop mixture. For example, one could plant corn at the rate of 50,000 plants per hectare in equally spaced rows one meter apart or in pairs of rows which are 0.25 meter apart and 1.75 meters between pairs. Then, one could intercrop the latter spatial arrangement with peas, beans, or some other annual legume at, say, 60,000 plants per hectare. The sole crop would have 50,000 corn plants per hectare while the intercrop mixture would have 50,000 corn plants plus 60,000 legume plants, or a total of 110,000 plants per hectare. Perhaps one should not be concerned about density, but should only consider total calories, profit, or some other joint function of yields. However, the above example illustrates another problem in designing intercrop experiments. If the legume sole crop was also at 60,000 plants per hectare in the above example, one wonders if the maximum yield of the intercrop wouldn't be obtained at a lower density for each crop, say, 35,000 corn plants and 45,000 legume plants per hectare.

5. Discussion

We have found the statistical design and analysis of intercrop experiments to be a challenging area of research. The statistician must become involved with much more than the mathematical niceties of Galois Field, Hilbert spaces, Borel sets, and power curves. He must learn something about agriculture and be creative in selecting models. Since it represents the most important unsolved statistical problem for tropical agriculture, and is applicable in many other areas, we find it a rewarding area of research and are preparing a manual on the topic.

Literature Cited

Eberhart, S. A. and Gardner, C. O. (1966). A general model for genetic effects. Biometrics 22, 864-881.

Federer, W. T. (1979). Statistical designs and response models for mixtures of cultivars. Agronomy Journal 71(5), 701-706.

Federer, W. T., Connigale, J.C., Rutger, J. N. and Wijesinha, A. (1981). Statistical analyses of yields from uniblends and biblends of eight dry bean cultivars. Crop Science (accepted for publication).

Hall, D. B. (1976). Mixing designs: A general model, simplifications, and some minimal designs. M.S. Thesis, Cornell University, May.

Mead, R. and Riley, J. (1981). A review of statistical ideas relevant to intercropping research. J. Royal Stat. Soc. A, 144(4), (Paper read on May 13, 1981).

Mead, R. and Stern, R. D. (1980). Designing experiments for intercropping research. Experimental Agriculture 16, 329-342.

Mead, R. and Willey, R. W. (1980). The concept of land equivalent ratio and advantages in yields from intercropping. Experimental Agriculture 16, 217-228.

Pearce, S. C. and Edmondson, S. C. (1981). Historical data as a guide to selecting farming systems with two crops. (Unpublished paper)

Pearce, S. C. and Gilliver, B. (1978). The statistical analysis of data from intercropping experiments. Journal of Agricultural Science 91, 625-632.

Pearce, S. C. and Gilliver, B. (1979). Graphical assessment of intercropping methods. Journal of Agricultural Science 93, 51-58.

Wijesinha, A., Federer, W. T., Carvalho, J. R. P., and Portes, T. de A. (1980). Some statistical analyses for a maize and beans intercropping experiment. BU-720-M in the Mimeo Series of the Biometrics Unit, Cornell University. (Submitted to Crop Science)