Statistical Analyses for Intercropping Experiments

by

Walter T. Federer
Cornell University
Ithaca, N.Y. 14853

BU-880-M May 1985

Abstract

Statistical methodology for analyzing intercropping experiments was developed over the last 20 years and is being developed at present. Considerably more research is required for the many and diverse types of experiments involving sole crops (crops grown alone) and mixtures of crops (intercrops) grown together or in sequence. The growing of two or more crops together or in sequence is known as intercropping. An outline of twenty chapters of a book on the statistical design and analysis of intercropping experiments is presented. A number of the statistical analyses in the book are briefly described. Sections 2 to 8 relate to analyses for two crops in a mixture along with sole crops. Sections 9 to 15 discuss analyses for three or more crops in a mixture in addition to sole crops and mixtures of two crops. It is stressed that it is dangerous to extrapolate from sole crop responses to mixtures of two crops and from mixtures of k crops to mixtures of k + 1 crops. Many of the data sets examined produced unexpected and sometimes surprising results. The last section discusses other areas of application, e.g., survey sampling, nutrition, education, medicine, and recreation, where these results can be utilized.

Statistical Analyses for Intercropping Experiments

by

Walter T. Federer
Cornell University
Ithaca, N.Y. 14853

BU-880-M*

May 1985

1. Introduction

Intercropping investigations involves the growing of two or more crops on the same area of land either simultaneously, partially at the same time, or sequentially. It is a centuries old practice in tropical agriculture, and to some extent in temperate zone agriculture. Agricultural, biological, and statistical investigations has tended to ignore the problems of research in this area. Statistical analysis; of intercropping investigations is considered to be the most important unsolved statistical question related to research in tropical agriculture. It is an area neglected by all except a handful of statisticians. A computer search of statistical literature resulted in the single paper citation for Mead and Riley (1981). This is an excellent paper, though limited in outlook for the broad range of statistical analyses useful in intercropping research.

---

* In the Technical Report Series of the Biometrics Unit.

To acquaint the statistical profession with relevant procedures and to fill a need by intercropping researchers, a book is being published by this author on the topic. The table of contents is:

It is necessary to fully comprehend the nature of two crop mixtures before proceeding to anything more difficult. The interpretational difficulty increases by an order in magnitude when going from sole crop (crops grown alone) experiments to experiments with sole crops and biblends (mixture of two crops.) It goes up another order in magnitude in going from intercropping experiments with two crops to experiments involving mixtures of three or more crops. In addition to the interpretational difficulty, it is dangerous to extrapolate from sole crops to biblends and from biblends to mixtures involving three or more crops. It is dangerous to extrapolate from lower densities to higher ones. Many, if not most, experiments contain an unexpected result.

A number of statistical analyses found useful for intercropping investigations are discussed below. The topics follow the table of contents of a forthcoming book that is outlined above.

2. One Main Crop Plus one Supplementary Crop

The experiment designs found useful for sole crops will be the same ones found useful for one main crop grown with a supplementary crop. The treatment design consists of the varieties of a main crop grown as sole crops and in combination with varieties of the supplementary crop. To illustrate, suppose that five = c_m varieties of maize are to be grown alone and in combination with six = c_b varieties of beans. A single density for maize and for beans is selected, i.e. plant population per hectare is not a variable. The treatment design would be:

Cropping System

Maize Variety	Sole	Bean Variety 1	2	3	4	5	6 = C_b
1							
2							
3							
4							
5 = c_m							

There would be $v = c_m + c_b c_m = 36$ treatments composed of five sole crops and 30 biblends. Experiment designs appropriate for 36 treatments would be used (see e.g., Federer and Kirton, 1984.)

Statistical analyses for experiments in a given experiment design and for the above treatment design would involve the same types of statistical analyses as used for sole crop experiments (see e.g., Snedecor and Cochran, 1967.) Some common statistical procedures used would be

(i) single (or subsets of) degree(s) of freedom contrasts,

(ii) multiple comparisons procedures,

(iii) subset selection procedures,

(iv) covariance analyses, and

(v) multivariate analyses.

Some additional statistical analyses found useful for yields are:

(vi) Tukey's one-degree-of-freedom analysis for the crop one by crop two interaction,

(vii) Finlay-Wilkinson (1963) analysis for mixtures,

(viii) tests for interaction given that one or more of the c_m maize varieties are standards for comparison, and

(ix) yields of main crop are not to be reduced by more than a fixed percentage.

3. Two Main Crops – Density Constant

Experiment design considerations for biblends when both crops are main crops, are the same as discussed in Section 2. The treatment design would have sole crops of both crops included; otherwise, it is the same as discussed in Section 2. Statistical analyses on the yields of each crop separately would follow that outlined in the previous section.

In order to evaluate cropping systems and to compare biblend production with sole crop production, it is necessary to combine the yields of both crops in some meaningful manner. An economic point of view would place a value, v_i, on the produce from crop i, say Y_i and use $V = v_1 Y_1 + v_2 Y_2$. If v_i are prices, it might be more realistic to use ratios of prices, which are more stable, and use relative values $V^* = Y_1 + Y_2(v_2/v_1)$. For sole crops, V(or V^* could be obtained by putting $Y_2 = 0$ for crop one and $Y_1 = 0$ for crop two. A nutritional point of view would convert the yield to calories and/or protein and use a measure of the form: $C = c_1 Y_1 + c_2 Y_2$, where c_i is a calorie (or protein) conversion factor. An agronomic or land use point of view would consider a linear combination of yields of the form:

$$L = \frac{Y_{b1}}{Y_{s1}} + \frac{Y_{b2}}{Y_{s2}} \quad ,$$

where Y_{bi} is the yield of crop i in a biblend mixture and Y_{si} is the yield of crop i grown as a sole crop. There are many forms of L_i, which is called relative yield or land equivalent ratio. The component yields of the mixture are put into proportions of yields obtained from sole crop yields. Since yields may vary considerably, a ratio of sole crop yields might be more stable. In this case a "relative land equivalent" ratio would be computed as

$$L^* = Y_{b1} + Y_{b2}\left(Y_{s1}/Y_{s2}\right) \quad .$$

A statistical point of view would use a discriminant function analysis and construct a canonical variable of the form:

$$D = Y_{b1} + bY_{b2} \ ,$$

where b is chosen to maximize the ratio, treatment sum of squares divided treatment plus error sums of squares.

The first three linear combinations given above, i.e., V, C, and L are readily interpretable quantities by a researcher or a farmer. The last one D is not and sole crop yields cannot be compared with D, but can be with V, C, and L. Although a statistician's first thoughts in combining yields most likely would be to use multivariate analyses, this would not be the correct thing to do as comparisons of sole crop yields and farming system yields cannot be made and the canonical variable has no practical meaning in the sense that C, V, and L do. Some aspects of multivariate analyses have been found useful by Pearce and Gilliver (1978, 1979) in studying the nature of response from mixtures.

Statistical analyses for linear combinations C, V, and L, are straightforward. Those outlined in the previous section may be utilized. These created functions of yield may be used in the same manner as canonical variables from a discriminant function analysis, i.e., univariate analyses are performed on the canonical variables. It is possible to combine value and land use by taking the ratio $Y_{s1}v_1/Y_{s2}v_2 = R$ and using the created function of yields $Y_1 + RY_2$. It does not appear realistic to combine variables other than yield variables as described above.

4. Two Main Crops — Density Variable

Plant populations per hectare in sole crops and in biblends need to be considered seriously in conducting intercropping investigations. Crop densities maximizing yields Y_i, or linear combinations of yield V, C, and D, are desired. Using univariate analyses, a multiple comparisons or subset selection procedure may be used to pick the "optimal" densities for the crops. A useful procedure would be to model yield as a function of plant density. Within narrow ranges of densities, a linear approximation of the form has been found to be useful:

$$Y_{ii\ell k} = \beta_{0i} + \rho_k + \beta_{1i} d_{i\ell_i} + \varepsilon_{ii\ell k} \quad ,$$

where $Y_{ii\ell k}$ is the yield of the ith crop as a sole crop, β_{0i} is an intercept, β_{1i} is a linear regression coefficient, $d_{i\ell_i}$ is the density ℓ_i for crop i, ρ_k is the effect of block k, and $\varepsilon_{ii\ell k}$ is a random error term with mean zero and variance σ^2_ε. Note that a variety of other functional relations could be used to model yield as a function of density. Using the above form, the yields of crop i in the mixture ij of two crops may be expressed as

$$Y_{i(j)\ell_1\ell_2k} = \beta_{0i} + \beta_{1i}d_{i\ell_i} + \gamma_{i(j)}\left(d_{i\ell_i}, d_{j\ell_j}\right) + \varepsilon_{i(j)\ell_1\ell_2k}$$

where $\gamma_{i(j)}(d_{i\ell_i}, d_{j\ell_j})$ is an additive effect on the yield of crop i due to its being intercropped with crop j at the corresponding densities $d_{i\ell_i}$ and $d_{j\ell_j}$. A large positive value of $\gamma_{i(j)}(d_{i\ell_i}, d_{j\ell_j})$ is desired. When there are many lines of a cultivar in an investigation, the above analysis may be conducted for each line. Then, analyses over all lines can then be obtained.

5. Modeling Responses for Sole Crops and Biblends - Two Responses

In many situations, responses for both components of a mixture are available. The crops may be intermingled but distinct in type so that responses for each crop are obtained, or the crops may be spatially separated and again responses for each crop are available. For treatment designs containing all sole crops and all possible combinations of lines of crops in mixtures of two, response model equations can be constructed which have measures of a general mixing ability (gma) effect and of a specific mixing ability (sma) effect of a line or crop. To illustrate, suppose that it was desired to compare yields of v = five bean cultivars as sole crops and in mixtures of two. The $v(v + 1)/2 = 15$ combinations would be:

Cultivar	1	2	3	4	5
1	S	B	B	B	B
2		S	B	B	B
3			S	B	B
4				S	B
5					S

where S stands for sole crop and B denotes a biblend. With such a treatment design in a randomized complete block design, one possible linear model is:

Sole crop i:

$$Y_{hii} = \mu + \rho_h + \tau_i + \varepsilon_{hii} ,$$

where $\mu + \rho_k$ is a block mean effect, τ_i is cultivar effect, and ε_{hii} have zero mean and common variance σ^2_ε.

Biblend ij:

$$Y_{hi(j)b} = \frac{1}{2} (\mu + \rho_h + \tau_i + \delta_i) + \gamma_{i(j)} + \varepsilon_{hi(j)b} ,$$

$$Y_{h(i)jb} = \frac{1}{2} (\mu + \rho_h + \tau_j + \delta_j) + \gamma_{(i)j} + \varepsilon_{h(i)jb}$$

where $Y_{hi(j)b}$ is the yield of cultivar i from the mixture ij, μ, ρ_h, and τ_i are as defined for sole crop, δ_i is a general combining ability effect for cultivar i when grown in biblends, γ_{ij} is an interaction effect for crop i in the presence of crop j, and the $\varepsilon_{hi(j)b}$ are error components for cultivar i responses which have zero mean and common variance $\sigma^2_\varepsilon/2$. The coefficient 1/2 is included in order to have the μ, ρ_h, τ_i, and δ_i from the biblends on the same basis as the corresponding <u>parameters</u> for sole crops. With two cultivars on the same area of land as the sole crops, each crop response can only contribute 1/2 to μ, ρ_h, and τ_i. Response model equations can easily be constructed for the case where one crop occupies a proportion p of the area and the second crop occupies 1 - p of the area. In this case, care must be taken in defining an interaction effect. An interaction is defined to relate to two items in equal proportions. To interact, both must be present. When $p < 1/2$, only 2p of the total material in an experimental unit is available to interact on a 1:1 basis;

1 - 2p of the material is not available. If some such definition as the above is taken, interaction effects will be invariant with respect to changing proportions p.

Note that when other treatment designs are used, other models can be constructed. For example, suppose that only a subset of the v(v - 1)/2 biblends were included in a experiment along with sole crops. The parameters μ, ρ_h, τ_i, and $\delta_i/2 + \gamma_{i(j)} = \gamma^*_{i(j)}$ can be estimated. It is not possible to obtain solutions for $\delta_i/2$ and $\gamma_{i(j)}$ but only their sum. If the experimenter were willing to assume that the $\gamma_{i(j)}$ not present were all zero, then solutions are possible. This is considered to be an unrealistic assumption.

6. Modeling Responses for Sole Crops and Biblends – One Response

For certain types of mixtures, such as, e.g., a diallel crossing experiment, it is impossible or difficult to obtain responses for both components of a biblend. Experiments involving sole crops and mixtures of two lines of a cultivar where the lines are not phenotypically distinct or are not spatially separated would be found for wheat, beans, and many other crops. In mixtures of grasses and legumes in hay it is difficult to obtain the separate responses for each member in the mixture. Several response models are available. For a randomized complete block design and the treatment design involving sole crops and all possible biblends, the following pair of equations for sole crop and biblend yields has been proposed (Federer *et al.*, 1982):

$$Y_{hiis} = \mu + \rho_h + \tau_i + \varepsilon_{hii}$$

$$Y_{hijb} = \mu + \rho_h + (\tau_i + \delta_i + \tau_j + \delta_j)/2 + \gamma_{ij} + \varepsilon_{hij}$$

where the effects are as defined in the previous section except for γ_{ij} which is an interaction component for specific mixing ability. Note that γ_{ij} is equal to the sum $\gamma_{i(j)} + \gamma_{(i)j}$. These last two components cannot be estimated unless individual responses are available whereas γ_{ij} can be estimated when only the combined response is available.

Another treatment design would be sole crops, all combinations, and all reciprocals. To illustrate, suppose that v = 5 wheat varieties are available, and the experimenter wishes to have all varieties bordered by every other variety and itself. Responses from border rows are not obtained. The v^2 = 25 treatments would be:

	Wheat Variety				
Border	1	2	3	4	5
1	S	B	B	B	B
2	B	S	B	B	B
3	B	B	S	B	B
4	B	B	B	S	B
5	B	B	B	B	S

where S denotes sole crop and B denotes the mixture. Note that variety 1 bordered by variety 2 is <u>not</u> the same as variety 2 bordered by variety 1. One set of response models for sole crop and biblends respectively is:

$$Y_{hii} = \mu + \rho_h + \tau_i + \varepsilon_{hii}$$

and

$$Y_{hijb} = \mu + \rho_h + \tau_i + \delta_i + \gamma_{ij} + \varepsilon_{hij,i \neq j},$$

where μ, ρ_h, τ_i, δ_i, ε_{hii}, and ε_{hij} are as defined as above and γ_{ij} is a within variety interaction term with $\gamma_{ii} = 0$; γ_{ij} is an interaction term for crop i when bordered by crop j.

A second response model equation for the above treatment design would be the one for a two-factor (crops and borders) factorial:

$$Y_{hij} = \mu + \rho_h + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{hij} \quad ,$$

where α_i is the effect of crop i, β_j is the effect of border j, and $\alpha\beta_{ij}$ is an interaction term. Such a model would not be too realistic in a variety of situations since sole crop responses may be quite different from biblend responses.

A third model is adapted from Martin (1980) and is the previous model with the following change:

$$\alpha\beta_{ij} = \eta_{ij} + \omega_{ij} + \kappa_{ij} \quad ,$$

where $\eta_{ij} = \eta$ for $i = j$ and $\eta_{ij} = -\eta/(v-1)$ for $i \neq j$, $\omega_{ij} = (\alpha\beta_{ij} + \alpha\beta_{ji})/2 + \eta/(v-1)$ for $i \neq j$, and $\kappa_{ij} = (\alpha\beta_{ij} - \alpha\beta_{ji})/2$.

A fourth model is a mixture of the previous ones and is

$$Y_{hijb} = \mu + \rho_h + \tau_i + \delta_i + \beta'_j + \omega'_{ij} + \kappa_{ij} + \varepsilon_{hijb} \, ,$$

where β'_j and ω'_{ij} are similar to the above β_j and ω_{ij} but are conditional on the fact that $\alpha\beta_{ii} = 0$; the remaining parameters are as defined above.

Other situations will lead to the construction of other response model equations. Appropriate models will need to be constructed for the particular conditions encountered in an investigation.

7. Spatial and Density Arrangements

Spatial arrangements and density levels are very important items to consider in intercropping investigations. By spatial arrangement, we mean the pattern used for plants in a given area of land. The plants could be in rows, in hills, or drilled. The number of plants per hectare could be varied over a wide range. The following five items need to be studied for any intercropping investigation:

(i) spatial arrangement of crop one,

(ii) spatial arrangement of crop two,

(iii) density of crop one,

(iv) density of crop two, and

(v) intimacy of the two crops.

By intimacy we mean the closeness of plants of the two crops. If plants of the two crops are randomly mingled in the same row, we say that they are 100% intimate. Plants of the two crops in separate rows would be less intimate. If the two crops were isolated far enough to eliminate any interaction, they have zero intimacy. To illustrate, suppose that density is not a variable but intimacy and spatial arrangement are. One plan could be to have two crops, say maize and beans, in the same row with rows one meter apart. A second plan could be to double the density within rows and double the distance between rows. The density per hectare and intimacy would be the same but spatial arrangement would be different. A third plan would be to alternate rows of the two crops. The intimacy would be less than in the first two plans. Another plan commonly used for maize and beans in Brazil is one row of maize and two rows of beans alternating as below (M = maize and B = beans):

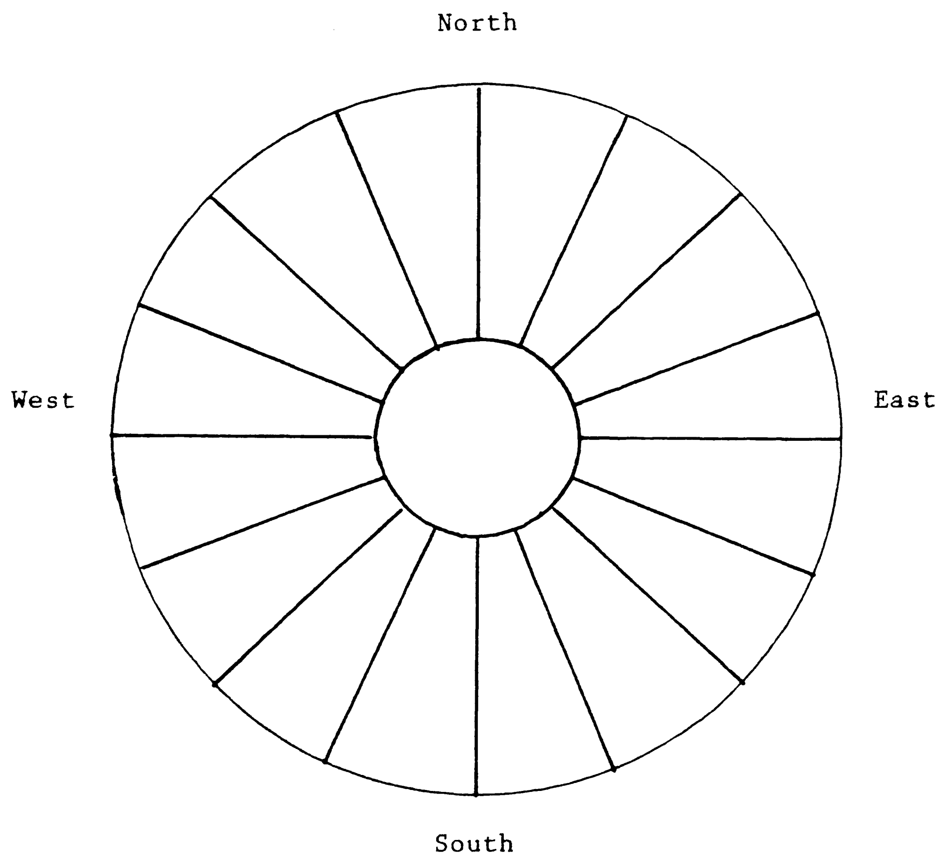
MBBMBBMBBMBBM··· .

The maize rows are one meter apart and the bean rows are one-half meter apart. A fifth plan would be:

MMBBBBMMBBBBMMBBBBMM··· .

The pairs of maize rows are 1.75 meters apart and the rows of a pair are 0.25 meter apart. The bean rows are one-half meter apart. The last plan could be the best as more light would be available for maize and for bean plants than in the previous plans. The rows should be oriented in a north-south direction in order to benefit from the additional light.

Several plans are available to study wide variations in density with a relatively small amount of material. They should be used to obtain information on ranges of density for future study. The best known of these is the fan design of Nelder (1962). There are several versions of this design. Another useful design has been suggested by B. N. Okigbo (1978). The design is a circle with orientation noted (see below). A small circle in the center is not used as some space is needed to start the rows. The row spacing becomes increasingly distant as one moves away from the center of the circle. The density within a row could be kept constant or the density per hectare could be kept constant by increasing the density within a row as one moves away from the center of a circle of the following nature:



The lines above could indicate the rows of plants. The above design could be for a single crop or for mixtures of two crops using the previously described plans for spatial arrangements and intimacy. A Nelder fan design would be one-quarter of the above and would be used if directional orientation were unimportant. Both the Okigbo-circle and the Nelder-fan designs are very parsimonious of space. One statistical analysis would be to divide the circle into concentric circles of equal areas. Yields would then be obtained for the areas of individual rows. The results could be plotted graphically to determine optimal yields or some regression function could be fitted to the yields. Optimal row distances and optimal densities for yield could then be obtained. These circles or fans could be constructed for various cropping systems and replicated over a range of conditions to be encountered in practice. It may be possible to determine optimal density, spatial arrangement, and intimacy well enough so that future experimentation is not necessary. However, it is likely that future experimentation will be needed to more precisely determine optimal values.

8. Variations and Additional Analyses

Many and diverse situations exist in intercropping research. One such area is to study the effect of replacing one crop in a mixture with a second crop with proportions ranging from zero to one. Given that p_a is the proportion for crop a and $1 - p_a = p_b$ is the proportion of crop b in the mixture and Y_{si} = yield of sole crop i, the computed value for a strictly replacement series would be $p_a Y_{sa} + p_b Y_{sb}$. If the yield of the mixture at proportion (p_a, p_b) was greater than this value, this would be termed cooperation. If less, then denote this as inhibition. If one crop is inhibited and the other exhibits cooperation, this would be denoted as compensation since the yield of one crop is increased and the other is decreased. For intercropping, proportions and crops showing a large amount of cooperation are desired.

Several other statistics have been developed for competition studies. A number of them are related to a land equivalent ratio.

Let $Y_{bi}/Y_{si} = L_i$ which is the proportional yield of the crop in a mixture relative to the crop grown alone. A land equivalent ratio is $L = L_1 + L_2$. A statistic was developed to compute *total effective area* for the case where A_i = area devoted to sole crop i and A_m = area devoted to the mixture of the two crops. Then, total effective area is computed as $A_1 + A_2 + LA_m$. A *relative crowding coefficient* is computed as $L_1 L_2/(1 - L_1)(1 - L_2)$. A *coefficient of aggressivity* to measure the dominance of one crop over another is computed as $L_1 - L_2$. A *competitive ratio index* is given by L_1/L_2. Each of these can be adjusted to the relative proportions $p_a:p_b$ of crop a and crop b in the mixture. Other coefficients have been suggested. A number relate to crop stability (an ill-defined term) and to "risk to farmers". Survival farming must take some form of these measures into account as a farmer needs to produce food every year in order to survive.

Another type of analysis suggested by B. R. Trenbath in his discussion of the Mead and Riley (1981) paper is linear programming. Here yields of the crops as sole crops and in mixtures is required. Then for a goal, say S units of starch and P units of protein, an optimal allocation of area to sole crops and to mixtures can be computed. A farmer can minimize land area needed to reach his primary goal (food production) and can use the remaining area of his farm for crops to achieve a secondary goal (say produce for sale). Economic studies make use of linear programming for some of their investigations.

9. One Main Crop with Two or More Supplementary Crops

Consideration of mixtures for more than two crops in the mixture would at first sight appear to be a straightforward extension of the procedures for two crops. This is not the case. To illustrate this for one main crop with supplementary crops, it would appear that one could simply follow the procedures described in Section 2, but consider the following treatment design and example. Barley was the main crop and only one barley variety was included in the experimental units along with barley in combinations of one cultivar plus barley, all possible combinations of three of the six cultivars with barley, and one combination of all six cultivars with barley. Plant numbers per experimental unit were kept constant and the same number of barley plants were harvested in every experimental unit. Barley as a sole crop was one of the treatments. In all there were 1 + 6 + 20 + 1 = 28 treatments. For a randomized complete block design and barley yields for variety g (one variety), one set of response equations is:

<u>Sole crop - variety g</u>

$$Y_{gh0} = \mu + \tau_g + \rho_h + \varepsilon_{gh} .$$

<u>Variety g plus one crop i</u>

$$Y_{ghi1} = \mu + \tau_g + \rho_h + \delta_i + \varepsilon_{ghi} .$$

<u>Variety g plus two crops i and j</u>

$$Y_{ghij2} = \mu + \tau_g + \rho_h + (\delta_i + \delta_j)/2 + \gamma_{ij} + \varepsilon_{ghij} .$$

<u>Variety g plus three crops i, j, and k</u>

$$Y_{ghijk3} = \mu + \tau_g + \rho_h + (\delta_i + \delta_j + \delta_k)/3 + 2(\gamma_{ij} + \gamma_{ik} + \gamma_{jk})/3 + \lambda_{ijk} + \varepsilon_{ghijk} .$$

$$\vdots$$

<u>Variety g plus all cultivars</u>

$$Y_{ghij\cdots v} = \mu + \tau_g + \rho_h + \delta_. + \gamma_. + \lambda_. + \cdots + \pi_{12\cdots v} + \varepsilon_{ghij\cdots} .$$

For the above example, mixtures of barley with two other crops were not included in the experiment. $\mu + \tau_g$ is the mean for barley variety g grown as a sole crop, ρ_h is the h'th block effect, δ_i is a general mixing effect of crop i on barley yields Y_{ghi1}, γ_{ij} is a bi-specific mixing effect of the combination of crops i and j on the yield of barley, λ_{ijk} is a tri-specific effect of the combination of crops i, j, and k on the yield of barley, $\pi_{12\cdots v}$ is a v-specific mixing effect of the combination of all v crops on the yield of barley, and all the εs are considered to have mean zero and common variance σ_ε^2. The assumption of common variance appears to be a realistic one for this experiment involving only barley yields.

10. Three or More Main Crops – Density Constant

A first step in analyzing data from an intercropping experiment containing mixtures of three or more crops is to obtain statistical analyses for each crop separately. The method of Section 9 may be used for this when appropriate. Response model equations for such experiments designed in a randomized complete blocks design, found useful are:

Sole crop g ($h = 1,2,\cdots,r;\ i = 1,2,\cdots,c_g$):

$$Y_{ghi} = \mu_g + \rho_{gh} + \tau_{gi} + \varepsilon_{ghi} .$$

Mixtures of three in proportions $p_1:p_2:p_3$, $p_1 \geq p_2 \geq p_3$):

Crop 1 yield, i'th line

$$Y_{1hi(jk)} = p_1(\mu_1 + \rho_{1h} + \tau_{1i} + \delta_{1i}) + 2p_2\gamma_{i(j)} + 2p_3\gamma_{i(k)} + 3p_3\pi_{i(jk)} + \varepsilon_{1hi(jk)} .$$

Crop 2 yield, j'th line

$$Y_{2h(i)j(k)} = p_2(\mu_2 + \rho_{2h} + \tau_{2j} + \delta_{2j}) + 2p_2\gamma_{(i)j} + 2p_3\gamma_{j(k)} + 3p_3\pi_{(i)j(k)} + \varepsilon_{2h(i)j(k)} .$$

Crop 3 yield, k'th line

$$Y_{3h(ij)k} = p_3(\mu_3 + \rho_{3h} + \tau_{3k} + \delta_{3k}) + 2p_3(\gamma_{(i)k} + \gamma_{(j)k}) + 3p_3\pi_{(ij)k} + \varepsilon_{3h(ij)k} ,$$

where interaction effects $\gamma_{i(j)}$, $\pi_{i(jk)}$, etc. are defined for equal amounts of material on an area basis, ε_{ghi} have zero mean and common variance

$\sigma^2_{g\varepsilon}$, $\varepsilon_{ghi(jk)}$ have zero mean and common variance $\sigma^2_{g\varepsilon 3} = \sigma^2_{g\varepsilon}p_g$, ρ_{gh} is a block effect for crop g, μ_g is a mean effect for crop g, and the subscripts in parentheses denote the other two crops in a mixture. Crops g, g*, and g' were taken to be 1, 2, and 3, respectively. The i'th line of crop g, the j'th line of crop g*, and the k'th line of crop g' is used. In experiments analyzed to date, only one line of each crop was included but the above equations are written to allow for one to c_g lines of each crop. Also, note that each crop's contribution to an interaction term can be estimated.

The construction of created variables as a linear combination of yields is straightforward from the two crop situation. For crop value, one uses $\Sigma_1^c v_g Y_g$ instead of $v_1Y_1 + v_2Y_2$. Or, all values v_g may be made proportional to a base crop value, say v_1; the created relative value will be $\Sigma_1^c Y_g(v_g/v_1)$. For calorie (or protein) value, the created variable $\Sigma_1^c c_g Y_g$ or $\Sigma_1^c Y_g(c_g/c_1)$ would be used. For land use values, the linear combination of yields $\Sigma_1^c Y_{gb}/Y_{gs} = \Sigma_1^c L_g$, or $\Sigma_1^c Y_{gb}(Y_{1s}/Y_{gs}) = Y_{1s}\Sigma_1^c L_i$ would be used for Y_{gb} = yield of crop g in a mixture and Y_{gs} = yield of crop g as a sole crop.

Multivariate discriminant function analyses are not usable (see Federer and Murty, 1984) for analyzing data from intercropping experiments. Multivariate theory needs considerable extension before it can be used. Problems of missing values, comparisons of sole crops with linear combinations of some of the crops, comparisons of different linear combinations, and the practical interpretation of the linear combination appear to make present concepts of multivariate theory unusable for intercropping data. Satisfying mathematical considerations and not practical interpretations is a vacuous solution for an experimenter trying to interpret results from an experiment.

11. Three or More Main Crops - Density Variable

With only two crops in a mixture, the assumption that the sole crop regression of yield on density holds for all densities of the second crop may be tenable in a small region of densities. With more than two crops in a mixture and with varying densities, this assumption may not be appropriate. To illustrate, consider mixtures of three crops gg^*g' for g, g^*, $g' = 1,\cdots,c$ crops at densities d_{ig}, d_{jg^*}, and $d_{kg'}$ for $i = 1,\cdots,c_g$, $j = 1,\cdots,c_{g^*}$, and $k = 1,\cdots,c_{g'}$. The regressions could be obtained for each of the $c_{g^*}c_{g'}$ density combinations and not just the sole crop. These regressions could be compared for homogeneity to ascertain whether the sole crop regression is appropriate for mixtures of three. If the regressions can be considered to be homogeneous or relatively so, the following response model equation for the yield of density combination $(d_{ig}, d_{jg^*}, d_{kg'})$ may be expressed as:

$$\begin{aligned} Y_{ghi(jk)}(d_{ig},d_{jg^*},d_{kg'}) &= \beta_{0g} + \rho_{gh} + \beta_{1g}d_{ig} \\ &+ \gamma_{i(jk)}(d_{ig},d_{jg^*},d_{kg'}) + \varepsilon_{ghi(jk)}(d_{ig},d_{jg^*},d_{kg'}) \end{aligned}$$

where $i = 1,\cdots,c_g$, $j = 1,\cdots,c_{g^*}$, and $k = 1,\cdots,c_{g'}$, β_{0g}, ρ_{gh}, and β_{1g} are as defined in Section 4, and $\varepsilon_{ghi(jk)}(d_{ig},d_{jg^*},d_{kg'})$ have zero mean and common variance $\sigma^2_{g\varepsilon}$. The $\gamma_{i(jk)}(d_{ig},d_{jg^*},d_{kg'})$ may be partitioned into an overall effect, an effect of crop g^* at density j, an effect of crop g' at density k, and an interaction effect for the jk'th densities of crops g^* and g'. These effects would relate to the yields of crop g.

12. Modeling Responses for Mixtures of Three or More Crops - Individual Crop Responses Available

Various response models for mixtures of two crops were discussed in Section 5. For mixtures of three of c cultivars, say i, j, and k, the following models are considered plausible for consideration using a RCBD:

Sole crop i

$$Y_{hi} = \mu + \tau_i + \rho_h + \varepsilon_{hi} .$$

Mixture ijk

Crop i yield =

$$Y_{hi(jk)} = (\mu + \rho_h + \tau_i + \delta_i)/3 + 2(\gamma_{i(j)} + \gamma_{i(k)})/3 + \pi_{i(jk)} + \varepsilon_{hi(jk)} .$$

Crop j yield =

$$Y_{h(i)j(k)} = (\mu + \rho_h + \tau_j + \delta_j)/3 + 2(\gamma_{(i)j} + \gamma_{j(k)})/3 + \pi_{(i)j(k)} + \varepsilon_{h(i)j(k)} .$$

Crop k yield =

$$Y_{h(ij)k} = (\mu + \rho_h + \tau_k + \delta_k)/3 + 2(\gamma_{(i)k} + \gamma_{(j)k})/3 + \pi_{(ij)k} + \varepsilon_{h(ij)k} .$$

A simpler form for crop i yield from a mixture of three would be

$$Y_{hi(jk)} = (\mu + \rho_h + \tau_i)/3 + \pi^*_{i(jk)} + \varepsilon_{hi(jk)}$$

where δ_i, $\gamma_{i(j)}$, $\gamma_{i(k)}$, and $\pi_{i(jk)}$ are all combined into an effect $\pi^*_{i(jk)}$. The interpretation of the parameters is the same as described in previous sections. Solutions for

$$\hat{\pi}^*_{i(\cdot\cdot)},\ \hat{\pi}^*_{i(j\cdot)},\ \hat{\pi}^*_{i(\cdot k)},\ \text{and}\ \hat{\pi}^*_{i(jk)},$$

subject to usual restrictions may be obtained when all possible combinations of crops are present. Otherwise, it is recommended that the above simpler form be used.

13. Modeling Responses for Mixtures of Three or More Crops - Individual Crop Responses Not Available

Suppose that sole crops and all possible combinations of three of the crops represent the treatments in a RCBD. Possible response model equations are:

Sole crop

$$Y_{hi} = \mu + \rho_h + \tau_i + \varepsilon_{hi} .$$

Mixture ijk

$$Y_{hijk} = \mu + \rho_h + (\tau_i + \delta_i + \tau_j + \delta_j + \tau_k + \delta_k)/3 + 2(\gamma_{ij} + \gamma_{ik} + \gamma_{jk})/3 + \pi_{ijk} + \varepsilon_{hijk} .$$

If all combinations were not present the model for mixtures may be simplified to:

$$Y_{hijk} = \mu + (\tau_i + \tau_j + \tau_k)/3 + \pi^*_{ijk} + \varepsilon_{hijk}$$

where a sum of general mixing (δ_i), bi-specific mixing (γ_{ij}), and tri-specific (π_{ijk}) effects would be represented in π^*_{ijk}.

Several other models described in Section 6 can be generalized to consider three or more crops in a mixture. When v^3 combinations of lines of three crops or factors are present, a three-factor factorial model may be used. Another response model for sole crops and mixtures of three crops i, j, and k would be:

Sole crop

$$Y_{his} = \mu + \rho_h + \tau_i + \varepsilon_{hi} .$$

Mixture ijk

$$Y_{hijkb} = \mu + \rho_h + \tau_i + \delta_i + \gamma_{ijk} + \varepsilon_{hijk}$$

where γ_{ijk} is an interaction effect within crop line i of component one of the mixture for lines j and k of the second and third components. Alternatively, γ_{ijk} could be an interaction effect within the combination ij. To illustrate, suppose that four lines of a crop, say A, B, C, D, are available, that center row yields only will be obtained, and that the center rows will be bordered on one or both sides by every line. For line A, the center and outside rows would be AAA, AAB, AAC, AAD, BAB, BAC, BAD, CAC, CAD, DAD. The interaction effects λ_{Ajk} would be the deviations of the quantities $\bar{y}_{\cdot ijkb} - \bar{y}_{\cdot i\cdot\cdot b}$, and the interaction effects λ_{ABk} would be the difference $\pm\ \bar{y}_{\cdot ABCb} - \bar{y}_{\cdot ABDb}$.

Martin (1980) states that his model does not extend to a three-factor factorial. A response model for a two-factor factorial in a RCBD would be:

$$Y_{hij} = \mu + \rho_h + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{hij} .$$

Martin's model deals with functions of the $\alpha\beta_{ij}$. A corresponding three-factor factorial response model would be:

$$Y_{hijk} = \mu + \rho_h + \alpha_i + \beta_j + \gamma_k + \alpha\beta_{ij} + \alpha\gamma_{ik} + \beta\gamma_{jk} + \alpha\beta\gamma_{ijk} + \varepsilon_{hijk} .$$

Construction of two-factor responses and using the previous model, $\alpha\beta_{ij}$, $\alpha\gamma_{ik}$, and $\beta\gamma_{jk}$ can all be partitioned. Partitioning of the three-factor interaction $\alpha\beta\gamma_{ijk}$ does not appear to be straightforward. One could collapse two of the factors into a single category and apply the previous Martin model. The other models discussed in Section 6 can likewise be extended.

14. Spatial, Density, and Intimacy Arrangements for Three or More Crops

For two crops, arrangements have been constructed to have one plant of crop one bordered by zero, one, two, three, and four plants of the second crop an equal number of times. Comparable plans for three or more crops have not been devised to date. As long as all plants of three or more cultivars (crops, lines of a crop, etc.) are randomly intermingled in an experimental unit, no difficulty arises. As soon as cultivars are placed in rows or planted in patterns, spatial patterns must be thoughtfully considered. The following items must be investigated for three crops:

(i) density of crop one,
(ii) density of crop two,
(iii) density of crop three,
(iv) spatial arrangement of crop one,
(v) spatial arrangement of crop two,
(vi) spatial arrangement of crop three,
(vii) intimacy of crops one and two,
(viii) intimacy of crops one and three, and
(ix) intimacy of crops two and three.

When using the Nelder fan or the Okigbo wheel, care must be taken in investigating orientation, density, spatial, and intimacy relations. These designs will be parsimonious of space and should be considered as obtaining preliminary results. More extensive investigation will more than likely be required in order to determine optimal conditions. The above considerations hold for mixtures of k of v crops.

15. Additional Statistics for Mixtures of Three or More Crops

Many of the statistics described in Section 8 may be extended to consider mixtures of three or more crops. The *total effective area* under three crops as sole crops, in mixtures of two, and in a mixture of three would be:

$$A_1 + A_2 + A_3 + L_{12}A_{m12} + L_{13}A_{m13} + L_{23}A_{m23} + L_{123}A_{m123} ,$$

where A_i = area under sole crop i, A_{mij} = area under mixture of two crops i and j, A_{m123} = area under the mixture of three crops, L_{ij} = land equivalent ratio for mixtures of crops i and j, and L_{123} is a land equivalent ratio for mixtures of the three crops.

A *coefficient of agressivity* for two crops in equal proportions of land area is $L_1 - L_2$. For three crops it would be $L_1 - (L_2 + L_3)/2$ for crop 1, $L_2 - (L_1 + L_3)/2$ for crop 2, and $L_3 - (L_1 + L_2)/2$ for crop 3. Extension to k crops is straightforward. L_i = yield of crop i mixture divided by yield of crop i as a sole crop.

A *competitive ratio index* for two crops in equal proportions of land area is L_1/L_2. For three crops, it would be $L_1/(L_2 + L_3)$, $L_2/(L_1 + L_3)$, and $L_3/(L_1 + L_2)$ for crops 1, 2, and 3, respectively.

A *relative crowding coefficient* for two crops is $L_1L_2/(1 - L_1)(1 - L_2)$. For k crops in a mixture, the coefficient would be $\Pi_1^k L_i/(1 - L_i)$.

Graphical representations for linear programming can be made for mixtures of two and three crops, but not for mixtures of four or more crops. However, linear programming techniques allow for k crops in a mixture and as sole crops.

16. Other Mixtures Where Statistical Techniques Are Useful

There are a large number of areas where the ideas and statistical procedures developed for intercropping can be used. For example, consider a survey sampling situation where answers are sought to sensitive, incriminating, and/or embarrassing questions. Direct questioning will not allow the surveyor to obtain this information. Anonymity of response is essential in order to obtain the information. Raghavarao and Federer (1979) have shown how to use the block total response procedure using supplemented and balanced incomplete block designs to obtain sensitive information. The respondent is required to give a total of answers to k of v questions. From the various block totals, estimates for the sample can be obtained without knowing individual responses. This is similar to knowing only the total response for a mixture rather than having the individual mixture component responses.

Other areas where these ideas can be utilized is in applications of drugs, therapies, medicines, recreational programs, physical training programs, educational programs, using sequences of courses and other mixtures, nutritional studies, use of pesticide and herbicide mixtures, and any other area where mixtures of components are involved. Studies in these areas to date have centered on mean comparisons of single or similar components, upon single responses for the mixture, and standard statistical procedures. Modeling aspects and competitive aspects have been ignored. Statistical theory has not provided adequate statistical methodology to do more than what is being done. It is a fruitful area for future research and application.

17. Literature Cited

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

Federer, W. T. and H. C. Kirton (1984). Incomplete block and lattice rectangle designs for v = 36 using F-square theory. No. BU-850-M in the Technical Report Series of the Biometrics Unit, Cornell University, September.

Federer, W. T. and B. R. Murty (1984). Uses and limitations of multivariate analyses in analysis of intercropping experiments. No. BU-858-M in the Technical Report Series of the Biometrics Unit, Cornell University, September.

Federer, W. T., J. C. Connigale, J. N. Rutger, and A. Wijesinha (1982). Statistical analyses of yields from uniblends and biblends of eight dry bean cultivars. *Crop Sci.* 22:111-114.

Finlay, K. W. and G. N. Wilkinson (1963). The analysis of adaptation in a plant-breeding programme. *Australian J. Agric. Res.* 14:742-754.

Martin, K. J. (1980). A partition of a two-factor interaction, with an agricultural example. *Applied Statistics* 29(2):149-155.

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

Nelder, J. A. (1962). New kinds of systematic designs for spacing experiments. *Biometrics* 18:283-307.

Okigbo, B. N. (1978). Personal communication. Intl. Inst. Tropical Agric., Ibadan, Nigeria.

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.

Raghavarao, D. and W. T. Federer (1979). Block total response as an alternative to the randomized response method in surveys. *J. Royal Statist. Soc., Series B* 41:40-45.

Snedecor, G. W. and W. G. Cochran (1967). *Statistical Methods*, 6th Edition. Iowa State University Press, Ames, Iowa.