EXPERIMENTAL STRATEGY FOR TRANSFERRING CROP PRODUCTION INFORMATION

by

BU-502-M Foster B. Cady June, 1974

Abstract

An increasingly important area of agricultural research is evaluating the transfer of experimental results from one region to another part of the world. Experiments need to be designed as an integrated series of experiments, with decisions made on controllable and other factors known to be important and measurable but beyond control of the experimenter. A transfer model which is site specific is formulated and procedures developed for estimating the model. The proposed strategy includes the selection of experimental sites over several years, determination of treatment and experimental designs, collection of both plot and site data, analysis of the data, and evaluation of the transfer hypothesis. A preliminary draft of this paper was presented at a workshop at the University of Hawaii, May 20-24, 1974 on Experimental Designs for Predicting Crop Productivity with Environmental and Economic Inputs, concerning two newly AID sponsored projects.

Paper No. BU-502-M in the Biometrics Unit Mimeo Series, Department of Plant Breeding and Biometry, Cornell University, Ithaca, New York 14850.

EXPERIMENTAL STRATEGY FOR TRANSFERRING CROP PRODUCTION INFORMATION

by

BU-502-M | Foster B. Cady | June, 1974

INTRODUCTION

One formulation of the UH/UPR/AID (University of Hawaii, University of Puerto Rico, Agency for International Development) soil fertility-crop production research projects begins with a statement of the population. The population includes units of upland soils in the tropics for closely related soil families as defined by the new U. S. Soil Taxonomy. The units in the population can be defined in terms of hectares, blocks, or in natural units of cultivated land. The units within the population exist by definition and need not be contiguous. Therefore a phase of a soil family, a family or closely related families, can be defined as a population even though units in the population might be on several continents. The soil family can then be a working definition of a population for the UH and UPR projects.

Clearly, experiments cannot be done on all units within the population and historically, soil fertility workers have used experiment station plots. The practice is to sample or select plots from the accessible subpopulation of station plots and to make recommendations to the entire population. Unfortunately, the experiment station plots sometimes are selected for special reasons.

Ideally, the areas chosen for experimentation are randomly selected from the units in the population. In the projects, random sampling will not be feasible primarily because a limited number of countries can be involved. Consequently, only part of the population is, in reality, available for selection of plots. It is implicitly assumed that inferences made to those units having zero probability of being selected for the experiments will not be affected.

The major hypothesis of the projects is that information from one set of experiments may be transferred to sites within the same soil family. The vehicle for transfer is a model estimated from the experimental data. For example, if applied phosphorus (P) is a factor of interest and experiments with one level are carried out, the calculated average response to the applied P is then used as a predictor for other sites in the same soil family. Usually more than one level is used and the estimated model is

$$\hat{Y} = b_0 + b_1P + b_2P^2 .$$

This estimated quadratic polynomial, where P is the level of applied P, is only one possibility for the transfer model. The square root is commonly used in fertility experimentation. Other possibilities are given in papers presented at the workshop by R. L. Anderson and B. G. Capó.

A major difficulty arises in using an estimated model for predicting the potential response to applied P at another location. Some factors known to influence yield, such as date of planting, can be held constant so that the estimated transfer function will perform well when used at another location, assuming the same planting date. However, other factors influencing yield cannot be controlled. Soil phosphorus will vary from one site to another within the same soil family due to past management of the surface horizon. Most climatic variables will also vary. Certain management variables can be held constant with sufficient resources, but probably not all. The effect of a variable which is specific to a site, and therefore called a site variable, is shown using soil phosphorus, p, as an example. Suppose two sites are used and the effect of the site variable is due only to the main effect. Then,

$$\hat{Y} = b_0 + b_1P + b_2P^2 + b_3p .$$

The estimated coefficient, b_3, can be estimated from soil phosphorus information at the two sites. In general, one numerical value will be available from each site for each site variable, and for each additional site parameter added to the model, an additional site has to be added in order to estimate the parameters of the model regardless of the number of plots at each site.

Suppose the measured soil phosphorus at the two sites is coded +1 and -1 for convenience. Then,

$$\hat{Y} = (b_0 + b_3) + b_1P + b_2P^2 \quad \text{for site 1}$$

and

$$\hat{Y} = (b_0 - b_3) + b_1P + b_2P^2 \quad \text{for site 2.}$$

Note that the curve representing the response to the applied phosphorus is the same for the two sites but the curves have different intercepts. It then follows that the economic optimum, found by equating $b_1 + 2b_2P$, the derivative of Y with respect to X, to the price-cost ratio is not affected by the soil phosphorus site variable.

However, the situation changes if the site variable interacts with the controlled variable; e.g., the difference between two levels of applied phosphorus depends on the level of soil phosphorus. The transfer model is

$$\hat{Y} = b_0 + b_1P + b_2P^2 + b_3p + b_4pP$$

and

$$\frac{\partial Y}{\partial P} = (b_1 + b_4p) + 2b_2P \ .$$

Now the economic optimum does depend on the level of soil phosphorus.

The response equation can also be written as

$$\hat{Y} = b_0 + (b_1 + b_4p)P + b_2P^2 + b_3p$$

showing that the estimated coefficient for the linear effect of P depends on the amount of soil phosphorus; i.e., the b_1 in the equation

$$\hat{Y} = b_0 + b_1 P + b_2 P^2 ,$$

calculated from data at each site, varies from one site to another. Later a method is described for testing the hypothesis that the linear coefficients are the same. If so, the response model, estimated from the data at all the sites, can be used for testing the transfer hypothesis. Invariably, however, the hypothesis of equal linear coefficients cannot be accepted. Now the transfer hypothesis has to be tested conditional upon the level of soil phosphorus. From a series of experiments, the estimated response model, including the coefficients b_3 and b_4, can be calculated. Then yield for additional (new) sites can be predicted using the estimated response model based on the series of experiments and p values from the new sites. Positive evidence for the transfer hypothesis is given by "good" agreement between the predicted yields and the actual yields from the new site.

One site variable, soil phosphorus, has been introduced. A large number of potential site variables exist and the approach described with only one site variable appears to be overly simplified. Indeed, soil experimenters are aware of a large number of interactions. However, the relative importance of these interactions when predicting yields for new sites could be overestimated. (Unfortunately, this statement isn't documented well in the literature but is the feeling of some soil fertility experimenters who have done extensive field experimentation. C. B. McCants also made this comment during the workshop.) It is also interesting to note that in the combined analysis of the Laird data, only drought had an important interaction with applied nitrogen when the prediction sum of squares was used in place of the residual sum of squares as a criterion for variable selection.

Using farmers' fields is not a new idea but the practice of planning a series of experiments on farmers' fields to obtain basic information on soil fertility relationships has been limited, sporadic, and sometimes gave disappointing results. In the past, limited resources of trained experimentalists, finances, and equipment have narrowed the soil fertility specialist's experimentation. Furthermore, procedures for measuring variables reflecting the productivity status of a given environment, e.g., a measurement of drought which is associated with yield, have not been available. Also data analysis and interpretation have been difficult when combining several sets of data. Fortunately, many of these limitations are now removed and with the world population pressures for reducing the time period between initiation of research and recommendations, strategy has to be developed for going directly to a relatively (as compared with the past) large number of experiments in a single stage of experimentation with perhaps a year or two of preliminary experiments.

The proposed strategy includes the selection of experimental sites over several years, determination of treatment and experimental designs, collection of both plot and site data, analysis of the data, and evaluation of the transfer hypothesis.

SELECTION OF EXPERIMENTAL SITES

If the units in the population are blocks of land of specified size, then, ideally, a random sample of blocks is selected and the experiments are carried out within the selected blocks. If the population is a geographically defined area, the randomization might include certain restrictions to insure a uniform distribution of experimental sites. Certain areas within the defined population might be more important than others and the sampling would be done in proportion to the weights of importance assigned to the various areas. Within these restrictions the basic procedure remains the same; namely, that random sampling procedures are used in the selection of experimental sites. In practice, the goal of random sampling without modification cannot be reached. A master list of all units in the population usually cannot be constructed easily. A substitute procedure is the use of a grid system across the area and a random sample of intersection points selected. For each selected point, for example, an experiment can be established in the grid north and east of the intersection point. Various schemes can be devised for selecting the particular site within the selected grids and again compromises will have to be made for practical considerations but with the retention of the spirit of random sampling.

Selection of sites has to be one of the most important stages of the experimentation when evaluating the transfer hypothesis. If sites are selected so that "identical pairs" are formed, then the hypothesis testing includes the ability for the experimenters to find sites in two locations which are very similar, at least for one growing season. On the other hand, a casual selection or a selection entirely dependent on factors such as the friendliness of the cooperator or closeness to a major highway, could seriously limit the possibility of supporting the transfer hypothesis. A great deal of preliminary work has to be accomplished before the final selection. One approach would be the selection of a relatively large number of potential sites using the previously described grid system. These potential sites are located and measurements made on those site variables where information can be ascertained before the experiments. The actual sites are then selected based on a number of practical considerations in addition to ensuring that a range of each of the measured site variables (also a two-dimensional scatter with a pair of site variables) is retained.

The number of sites and replications at each site have to be determined. For reasons more fully developed in a later section, the desired number of sites is probably larger than the resources available. Briefly, certain soil variables and most environmental and management variables are measurable only at a site level. For example, even though multiple measurements are made on drought effects, and careful rainfall and soil moisture records are kept, a single valued index is calculated for an individual site for a given year. For response data and certain soil variables, however, observations are measured each year on all plots at an individual site. For each site variable used in the combined analysis of data, it is important that a range of values for each site variable exist. If the range does not exist in the population, then that variable should not be used. Additional restrictions on the site selecting procedure are sometimes needed to insure a sufficient range in the sample for each of the site variables. The site variables are used in the analysis to interpret the expected interaction between sites and treatments. The number of locations has to equal at least the number of site variables and hopefully the number of locations will be much larger than the number of site variables to allow for testing the adequacy of the site variables in explaining the treatment by location interaction. Twenty locations per year would not be a large number for many soil fertility experiments. With a relatively large number of locations per year, the number of replications per location can be minimized unless the analysis of data at each site is important. Otherwise two replications per location is sufficient for detection of gross plot errors, for obtaining standard errors and coefficients of variation at each location, and for providing sufficient degrees of freedom for the pooled error term in the combined analysis of the data.

Ideally, sufficient years are necessary so that a range and distribution of climatic conditions covered during the experimental stage approximates the long-term range and distribution. Again compromise with resources has to be made. In any case an analysis and interpretation is made after each year of experimentation and predictions made for the forthcoming year. With this type of sequential approach the magnitude of the prediction errors can be considered each year and, together with the range of climatic conditions experienced, a decision can be made concerning the additional number of years needed. As a general statement, a minimum of three or four years seems necessary.

DETERMINATION OF TREATMENT AND EXPERIMENTAL DESIGNS

The treatments are combinations of factors known to be most important in increasing yields and are inputs which are under the control of farmers. The number of factors is usually held to two or three and frequently are major nutrients and/or management factors such as date of planting. Management factors can be difficult to define and handle in a series of experiments over a large geographical area. These combinations of controllable factors usually consist of complete or incomplete factorials. Response surface designs can be considered as special cases of incomplete factorials. The particular set of combinations used in a series of experiments depend on several considerations. A practical need exists for plotting the data at each site soon after the yields are available and for preliminary examination of the data before any combined analysis is attempted. The treatment combinations also depend on variance and bias of various statistical estimators, e.g., regression coefficients and yield predictions. A more detailed discussion of these considerations was given by R. J. Laird at the workshop and his manuscript covers details of treatment design. A decision on the overall range of the controlled factors is undoubtedly more critical than generally assumed.

The experimental design can usually be a randomized complete block design. The additional cost of having blocks instead of a completely randomized design is minimal, while the experimental error reduction could be considerable. Trying to improve the precision of the individual experiments through experimental design is a secondary consideration in a series of experiments where the focus is the analysis and interpretation of data combined over the environments. Complications with unexpected happenings, such as a missing plot, usually would preclude use of incomplete block designs.

COLLECTION OF PLOT AND SITE DATA

Soil fertility experimentation has a history of collecting plot data, but experience in measuring site variables is much less extensive. Site variables are factors believed to affect crop yields but cannot be controlled at a constant level across the experiments. Neither is it desired to control these site variables at a constant level, since the effect of site variables on the relationship between yield and the controlled factors is one of the major reasons for the series of experiments. Recently, extensive work has been done measuring climatic variables under field conditions. Equipment is becoming available and numerical work on correlations between plant symptoms in the field, as indicators of drought and disease, and yield is now receiving increased attention.

ANALYSIS OF DATA

The data include one or more response variables measured from the experimental plots where a particular combination of controlled fertility or management factors had been applied. In addition, data are collected on a number of uncontrolled site factors including soil, climatic, and management variables.

For a given crop, yield is a function of soil, environmental, and management factors. Conceptually, this function can be written as a statistical model but several problems arise. There is no general agreement on the true functional form expressing the relationship between yield and an input variable. Polynomials and various nonlinear (in the parameters) functional forms have been tested over the years with no clear-cut indication that one form is overall best. Consequently, at least for convenience, a linear functional form, such as the quadratic polynomial or the square root model, is more generally chosen. Discontinuous models could also be considered. A second problem is a subject matter problem in that all the factors for each of the soil, climatic, and management classes are not known. If all the variables for each class could be listed from basic subject matter considerations, it would not be known if all are sufficiently important to be included in the functional model. Or, if a given variable is important in general, it might not be important in the defined population for several reasons:

(i) A sufficient range of values for the variable might not exist in the population.

(ii) A sufficient range of values exist but most of the range is beyond the "critical level" and consequently variability in yield cannot be explained by that variable.

(iii) The variable cannot be measured sufficiently well under field conditions, at least relative to the magnitude of the effect.

(iv) The variable is important but is suppressed by another variable especially within a given year.

(v) The variable has an important interaction effect with another input variable but in the sample of sites, low values of one variable are associated with low values of the other variable or high with high, but the low-high and high-low combinations don't appear in the data.

Consequently, in the model-building problem the data are used first for making tentative decisions as to which of the measured variables should be included, and their functional form. Then the parameters of the model are estimated with the same data.

The first step in the data analysis should be an analysis of the individual experiments. Suppose a two-factor fertility experiment with a partial factorial of 13 nitrogen and phosphorus treatment combinations in a randomized complete block design with two replications was used at each location. Analysis of variance calculations would give the usual sum of squares for blocks, treatments, and experimental error. The regression of yield on nitrogen and phosphorus, including the intercept term, the linear and quadratic effects, and the linear by linear interaction will give a fitted model sum of squares. The difference between the treatment sum of squares and the fitted model sum of squares is the lack of fit which can be tested against the experimental error mean square for adequacy of the fitted model. Other checks include examination of the signs and magnitudes of the estimated coefficients, and the magnitude of the calculated coefficient of variation.

The regression model could be fitted using the yield data from all the experiments. Experience has shown that the overall model with applied nitrogen and phosphorus terms will give a poor fit, using R^2 as a criterion for goodness of the fitted model. This is not surprising since the uncontrolled variables are affecting the response to fertilizer at each location differentially. For example, the levels of soil nitrogen and phosphorus can be expected to affect the response to the applied quantities. Consequently, a major objective in the interpretation of data from a series of experiments is the determination of a general yield model or equation with the inclusion of the uncontrolled variables. The functional relationship between yield and the applied fertilizer variables as influenced by the site variables can be calculated and used in estimating fertilizer needs for specific conditions.

Historically, the general yield equation has been estimated by least squares using linear, quadratic, and cross product terms of the measured independent variables. Again the results usually are not satisfactory. The sign or magnitude of the estimated coefficients will not be consistent with known information and an unexpected high proportion of the estimated coefficients will be smaller than their estimated standard errors. Even worse will be the performance of these estimated equations when used for predictions with different but similar sets of data.

The number of terms in the estimated equation is then usually reduced using a sequential procedure where variables are selected if the entering variables statistically lowers the residual sum of squares. Improvements through the stepwise regression procedure can be noticed with the selected reduced models, but most of the basic problems remain and the additional problem of certain supposedly important variables not appearing in the selected reduced model occurs.

It is known that introducing an additional variable into a yield equation cannot decrease the variance of a predicted observation; in fact, the variance will be increased in practical situations. However, failure to include a variable may result in a biased predicted observation. An ideal procedure would select variables which are important in reducing bias without selecting those which would unnecessarily add to the variance of a predicted value.

A new procedure has been developed for selecting predictor variables from a large number of potential ones. The procedure uses as a criterion the prediction sum of squares (PRESS) defined as

$$\sum_{i=1}^{n} (Y_i - \hat{Y}_{(i)})^2 ,$$

i.e., the sum of squares between Y_i, the observed response, and $\hat{Y}_{(i)}$, the predicted response, using a prediction equation where coefficients have been estimated excluding the i^{th} observation. The squared deviations are then added over the n observations. A sequential procedure is used to determine the order that variables will enter the yield equation. The first variable will be that one which has the smallest value of PRESS. It appears that for each of the p potential predictor variables, n regression equations would be calculated, the omitted observation predicted, and the n squared deviations added. The pn regressions are not calculated in practice due to an algorithm developed by Dr. David Allen. After the first predictor variable is selected, the procedure is repeated among the remaining p - 1 potential variables to find that variable which will give the smallest value of PRESS at the second stage.

One problem with using stepwise regression is that the criterion, the residual sum of squares, is reduced with each additional variable and a completely arbitrary decision is made for the cutoff point. In practice, the entering variable is usually tested at a significance level of .05 or .10. PRESS has a distinct advantage in that the criterion, the prediction sum of squares, decreases with entering variables to a minimum and then increases. The cutoff point can then be at the minimum or perhaps before, depending on the nature of the decreasing prediction sum of squares function. The main advantage of PRESS is that it should give better predictions of future responses for new combinations of the independent variable values. Details of this criterion and an example of its use are given in papers listed in the bibliography.

The problem encountered in the described variable selection procedures can be traced in a general sense to multicollinearity resulting from existing correlations among the uncontrolled variables and induced correlations from including quadratic and cross product terms. This problem can be avoided in part by preliminary tests on each of the uncontrolled factors before inclusion into a general yield equation. Returning to the individual site analyses, it can be seen that the estimated regression coefficients in the function between yield and the applied fertility variables are summary statistics or condensations of the original plot data. These calculated random variables can be used in preliminary tests of significance for information on the inclusion of certain site variables in the overall yield equation as main effects or as interactions with the applied fertility variables. Details of this approach will now be given.

At each site a model relating yield to the applied fertility variables will be fitted by least squares. For example, if the treatments were combinations of applied nitrogen and phosphorus and a quadratic polynomial is fitted to the yields, the fitted model for site t would be

$$\hat{Y}_t = b_{0t} + b_{1t}N + b_{2t}P + b_{3t}N^2 + b_{4t}P^2 + b_{5t}NP .$$

The six estimated regression coefficients attempt to explain the non-random variability at site t. If there were no other factors affecting yield, the same response would be observed at each site. For example, if β_{0t}, the parameter, is the same for all sites, i.e., $\beta_{0t} = \beta_0$ for all values of t, t = 1, 2, $\cdots$, s sites, then all the b_{0t} are estimating the same parameter and all the uncontrolled factors have no effect on the mean yield at site t, i.e., the main effect of all the uncontrolled variables is zero. This hypothesis can be tested with an F test:

$$F = \frac{\sum_{t=1}^{s} (b_{0t} - \bar{b}_0)^2 / s - 1}{c_{00} s^2}$$

where $\bar{b}_0$ is the mean of the b_{0t}, s^2 is the estimated experimental error from pooling the individual site experimental errors, and c_{00} is the first diagonal element from the inverse of the matrix of sum of squares and cross products formed from the applied fertility variables. If the same treatment design was used at each site, i.e., the same treatment combinations, then the inverse will be the same and c_{00} is a constant.

The same argument holds for all β_{it}, $i = 0, 2, \cdots, 5$. For example, if all the b_1 are estimating a single β_1 the linear effect of applied nitrogen is the same at all sites and no interaction exists between applied nitrogen (linear) and the uncontrolled site variables.

F tests for the homogeneity of the mean and linear terms will usually be significant, and the next step is to identify those measured sites or uncontrolled variables associated with the significant F values. The b_{it} values will now be used as the dependent variable. The correlations between the b_{it} and the measured site variables and basic soil fertility knowledge can be used to indicate those site variables that undoubtedly are important. At this stage each of the b_{it}, for example the b_{0t} values, can be regressed on the site variables

$$b_{0t} = \alpha_{00} + \alpha_{01}X_{1t} + \cdots + \alpha_{0p}X_{pt} + e_{0t}$$

where the α_{0j}, $j = 0, 1, \cdots, p$, are the parameters relating the estimated site means to the p uncontrolled variables and e_{0j} is a random error component.

Estimating the α_{0j} has some of the same problems as previously mentioned, namely that of multicollinearity among the X's. Two kinds of X usually can be identified:

(i) factors with sufficient ranges to almost always affect the regression coefficients, and

(ii) factors with ranges that might or might not affect the regression coefficients, depending on the sampling in the given series of experiments.

Factors of type (i) should definitely remain in the model and the parameters estimated by least squares. With factors of type (ii) a criterion such as the residual sum of squares (or a function of the residual sum of squares such as R^2 or the C_p statistic) or the prediction sum of squares (PRESS) should be used for a decision on inclusion or exclusion. For the included variables, the least squares procedure can be used for estimation or a biased estimation procedure such as ridge regression should be considered.

These multiple regressions are run for each of the coefficients where a significant F was found in an earlier step and the regressions are then substituted for the b_i in the original equation relating yield to the applied fertility variables to give a general yield equation.

EVALUATION OF THE TRANSFER HYPOTHESIS

Two data analysis procedures were outlined in the previous sections for selecting the variables to be included in a general yield equation and for estimating their parameters. The estimated final yield equation only has value if the estimated parameters can be used for new sets of input variable values (new values of the selected independent variables) which are similar to those in finding the estimated equation.

As stated in the introduction, the estimated response function for the applied nutrients can be transferred to another location if all the other factors are at the same level. Within the same soil family certain soil properties and long-term climatic factors are the same. Consequently, information from one country hopefully could be transferred, at least qualitatively, to another country. However, the transfer vehicle is a model estimated from yield data. Yields not only depend on the applied treatments but also on factors not constant within a soil family, e.g., management variables. In addition, the variation among sites within a country, for a variable supposedly constant within a soil family, can be important quantitatively and annual variation in climatic variables can be relatively large. These important site variables have to be included in a general yield equation or transfer model. A procedure for evaluating the transfer hypothesis is now formulated.

Let $Y_{ijk\ell}$ be the yield from the i^{th} country, the j^{th} site within the country, the k^{th} treatment or controlled variable, and the ℓ^{th} replication. For convenience, the following discussion will include two countries, i = 1, 2, s sites within each of the two countries, j = 1, 2, ···, s, and the same treatment design at each site, k = 1, 2, ···, t. Treatment means will be used in the analysis and the fourth subscript deleted. Also for convenience, applied phosphorus will be the only controlled variable and soil phosphorus the only site variable.

From the Y_{11k}, Y_{12k}, ···, Y_{1sk} means within country one, an estimated model can be calculated as

$$\hat{Y}_1 = a_1 + b_1P + c_1P^2 + d_1p + e_1Pp$$

where a_1, b_1, c_1, d_1, and e_1 are the partial regression coefficients estimated from country one data. Similarly, the estimated model for country two is

$$\hat{Y}_2 = a_2 + b_2P + c_2P^2 + d_2p + e_2Pp \; .$$

Note that the values of P are the t applied phosphorus levels which are constant at all sites and the values of p are the s soil phosphorus levels within each country, i.e., one value of p exists for each site, as measured from a composite soil sample. Also note that both estimated models can be written as

$$\hat{Y}_i = (a_i + d_ip) + (b_i + e_ip)P + c_iP^2 \; .$$

By knowing the numerical values of the estimated coefficients and assuming that no other factors influence yield, this model can hopefully be transferred to another site in the same or different country and do equally well in predicting yield using the measured soil phosphorus level at the new site.

Additional notation is needed. Let $\hat{Y}_{i(i)}$ be the predicted mean yields when the transfer model estimated from one country is used to predict yields for sites in the same country, e.g., $\hat{Y}_{1(1)}$ and $\hat{Y}_{2(2)}$. $\hat{Y}_{i(i')}$, where $i \neq i'$, denotes the predicted mean yields using a transfer model with the regression coefficients estimated from one country predicting yields for the other country using the soil phosphorus values for the latter country, e.g., $\hat{Y}_{1(2)}$ and $\hat{Y}_{2(1)}$. In all four cases t treatment means are predicted for each site. The values of P will vary, depending on the level of the controlled factor, while the value of p will be constant for a given site but vary from site to site.

If the estimated model can be transferred to the other country, then intuitively the $\hat{Y}_{1(2)}$ and $\hat{Y}_{2(2)}$ prediction models should perform equally well. Similarly, $\hat{Y}_{1(1)}$ and $\hat{Y}_{2(1)}$ should do equally well. A measure of discrepancy is the squared deviation between the observed mean yield and the predicted mean yield. These squared deviations can be summed over the treatments at a site and over the sites within a country. Following this reasoning, the expected value of the following ratio would be one if the transfer models were doing equally well predicting within the country as across countries.

$$\frac{\sum\limits_{jk}\sum(Y_{1jk} - \hat{Y}_{2(1)})^2 + \sum\limits_{jk}\sum(Y_{2jk} - \hat{Y}_{1(2)})^2}{\sum\limits_{jk}\sum(Y_{1jk} - \hat{Y}_{1(1)})^2 + \sum\limits_{jk}\sum(Y_{2jk} - \hat{Y}_{2(2)})^2} .$$

If the ratio is much larger than one, the transfer hypothesis would have to be rejected. If near one, the transfer hypothesis is supported in the sense that prediction is no worse going across countries than it is within countries. A ratio value of one does not necessarily support the hypothesis that response information can be transferred from one site to another within a country. The theoretical properties of the ratio are not known so that a test statistic could be used in the evaluation. Additional work is needed in this area of evaluation of the transfer hypothesis.

BIBLIOGRAPHY

[1] Allen, David M. (1971). The prediction sum of squares as a criterion for selecting predictor variables. University of Kentucky, Department of Statistics, Technical Report No. 23.

[2] Anderson, R. L., D. M. Allen, and F. B. Cady (1972). Selection of predictor variables in linear multiple regression. Chapter 1, _Statistical Papers in Honor of George W. Snedecor_. Iowa State University Press.

[3] Cady, Foster B. and David M. Allen (1972). Combining experiments to predict future yield data. _Agron. J._ _64_:211-214.

[4] Laird, R. J. and F. B. Cady (1969). Combined analysis of yield data from fertilizer experiments. _Agron. J._ _61_:829-834.