Conference on Applied Statistics in Agriculture 1993 - 5th Annual Conference Proceedings

# DESIGNING ALFALFA YIELD TRIALS FOR COMPARING LONG-TERM YIELDS

G. B. Schaalje

S. N. Acharya

# DESIGNING ALFALFA YIELD TRIALS FOR COMPARING LONG-TERM YIELDS

G. B. Schaalje and S. N. Acharya[1]

Department of Statistics, Brigham Young University, Provo, Utah 84602
[1]Research Station, Agriculture Canada, Lethbridge, Alberta Canada T1J 4B1

**Abstract.** An aspect of experimental design that must be taken into consideration for variety trials of perennial crops is the number of years to continue the trial. By tradition, alfalfa forage yield trials are harvested for three or four production years, but the consumers of information from these trials, the producers, often keep their stands in production for more than four years. This study developed a statistical efficiency measure for evaluating the adequacy of forage trial designs with specified numbers of years and replicates, based on a multivariate linear model. The measure was applied to data from four long-term trials grown in western Canada. Variances and covariances for varieties and for residual errors varied from trial to trial. Estimates of variety differences for the four-, five-, and six-year total yields using equally weighted combinations of three, four, and five years' data, respectively, were reasonably efficient. Four replicates were sufficient for the four- or five-year total yield, but more replicates were needed for efficient evaluation of the six-year total yield.

**Keywords.** experimental design, perennial crop, efficiency, multivariate linear model

## 1. Introduction

Alfalfa is an important perennial crop in western Canada, and for many years yield trials have been carried out with the purpose of selecting varieties with high 'long-term' forage yields. Most work on the design and analysis of yield trials has dealt with annual crops; little thought has been given to the special design problems associated with perennial crops, such as the number of years to continue the trials. This is a particularly tough question for alfalfa breeders because there is a discrepancy between how long producers keep their alfalfa fields in production (5 or more years), and how long agronomists recommend that alfalfa fields should be kept in production (3 or 4 years). A recent economic analysis agreed with producers that the optimal replacement time for alfalfa stands is longer than 4 years (Stauber and Goodman 1986). Even so, the question remains as to how long to continue alfalfa yield trials because it would be a waste of resources to continue such trials for 5 or 6 years if good estimates of the long-term yields could be obtained from 3 or 4 years' data. A literature search failed to produce any guidance on this question.

The objective of the present study is to develop statistical methods for determining the number of years and the number of replicates needed for efficient assessment of long-term yield differences among alfalfa varieties. It must be recognized, however, that there is no single solution to this question because important factors vary from trial to trial. Such factors include the particular entries used in the trial, weather patterns during the trial, and spatial and temporal variation associated with experimental plots used in the trial. Our approach in this paper is to develop general statistical methods and then apply these methods retrospectively to data from several longer-term trials to determine how many years and replicates would have been needed to adequately assess varietal differences in those specific cases, and to see if common patterns emerge from all data sets. This parallels the usual practice of field experimentation in which guidelines as to optimal plot and block sizes, for example, are gradually developed with experience over years (Pearce 1976). The four data sets available for this analysis were from alfalfa variety trials in western Canada which were harvested for five or six years.

## 2. Materials and Methods

### 2.1. Yield Trial Data

Yield data from Uniform Alfalfa Trials established in 1983, 1984, and 1985 at Lethbridge, Alberta and in 1985 at Swift Current, Saskatchewan, denoted by L83, L84, L85, and S85, respectively, were used. Forage yields were observed for five years on L83 and L84 and six years on L85 and S85. All trials had 20-25 varieties, were laid out in randomized complete blocks, and had no missing data. The three Lethbridge trials each had four replicates while the Swift Current trial had six replicates. All trials were grown under sprinkler irrigation. The number of cuts varied from year to year; the yearly forage yield from a plot was taken as the total dry matter harvested from all cuts of that plot during the year in question. Except for control varieties, the trials involved different sets of varieties; for the purposes of this paper, information on the varieties used in these trials is not important.

### 2.2. Statistical Model

Statistical developments in this paper are based on the following multivariate linear model for the five- or six-year series of yields from a given plot of a trial:

$$\mathbf{y}_{ij} = \mu + \mathbf{v}_i + \mathbf{b}_j + \mathbf{e}_{ij} \qquad (1)$$

where $\mathbf{y}_{ij}$ is a vector of the five or six observed yearly yields from the plot containing variety i in replicate j of the experiment, $\mu$ is a vector of yearly mean yields over all varieties and blocks, $\mathbf{v}_i$ is a vector of yearly deviations for entry i from the overall yearly means, $\mathbf{b}_j$ is a vector of yearly deviations from the yearly means for block j, and $\mathbf{e}_{ij}$ is a vector of yearly residuals associated with variety i in block j. The $\mathbf{e}_{ij}$ are assumed to be independent random vectors with mean vector $\mathbf{0}$ and variance-covariance matrix $\Sigma_\mathbf{e}$. Nothing is assumed about the structure of $\Sigma_\mathbf{e}$; the plot residuals can have different variances in different years, and can be correlated to varying degrees from year to year. For the purposes of this paper, it is also assumed that the variety deviations ($\mathbf{v}_i$) are random vectors, distributed with mean vector $\mathbf{0}$ and variance-covariance matrix $\Sigma_\mathbf{v}$, where again no assumptions are made about the structure of $\Sigma_\mathbf{v}$. Although it may seem odd to treat the variety effects in the yield trials as a random rather than a fixed factor, it is appropriate in the present context because the objective was to make inferences applicable to future yield trials involving unspecified alfalfa varieties. In a related context, Stroup and Mulitze (1991) pointed out that when variety effects are taken as a random factor, predictors of mean variety yields can be more efficient than those obtained taking varieties as a fixed factor.

This model was fitted to the yield trial data sets using multivariate analysis of variance via the SAS procedure GLM (SAS Institute, 1989). An estimate of $\Sigma_\mathbf{e}$ for each of the alfalfa yield trials was obtained directly from the output. An estimate of $\Sigma_\mathbf{v}$ for each of the trials was obtained using the method of moments. The estimate of the 'between variety' covariance matrix was set equal to its expected value in order to solve for $\Sigma_\mathbf{v}$. In the remainder of the study, the estimates of $\Sigma_\mathbf{v}$ and $\Sigma_\mathbf{e}$ are treated as if they are true values as is common when planning experiments.

A problem was encountered with the method of moments estimates for $\Sigma_\mathbf{v}$. The estimates were sometimes invalid in the sense that correlation coefficients, obtained by dividing estimated covariances by appropriate estimated standard deviations, were sometimes greater than 1. For this paper the problem was ignored because the correlations in question were always fairly close to 1; if adjustments were to be attempted, it was not clear whether to adjust the estimated covariances, variances, or both; and subsequent calculations did not depend on the correlations being valid. Future work on this problem is planned.

### 2.3. Notation for Predictors

The true long-term yield to be used for selecting varieties can be expressed as a linear combination of the expected five- or six-year yields, denoted by the vector product

$$\mathbf{m}' \,(\mu + \mathbf{v}_i) \qquad (2)$$

where $\mathbf{m}'$ is a vector of constants. For example, if there is interest in the differences among varieties for the total yield over the first five years of the trial,

$$\mathbf{m}' = (1\ 1\ 1\ 1\ 1\ 0).$$

Linear combinations of observed yields from the first three or four years of a trial are usually implicitly used as predictors of long-term yields because only short-term data (three or four years) are available. A short-term predictor of a specified long-term yield is denoted by the vector product

$$\mathbf{a}'\ \mathbf{y}_{i.} \tag{3}$$

where $\mathbf{y}_{i.}$ is the mean vector of the observed yearly yields for variety i over the replicates. For example, if the equally weighted average of the first three years' yields were to be used in predicting differences among varieties for some long-term yield,

$$\mathbf{a}' = (1/3\ 1/3\ 1/3\ 0\ 0\ 0).$$

The idea of this study is to develop criteria for evaluating and comparing different short-term predictors of specified long-term yields, involving different numbers of years and different weights for the yearly yields.

## 2.4. Assessing Predictability

The selection of the best varieties from a group of candidates implies that interest is in the differences among the varietal means rather than the varietal means themselves. Thus a reasonable criterion for evaluating a possible short-term predictor of a specified long-term yield is the expected error of the predicted difference in long-term yield between any two varieties, averaged over all possible pairs of varieties in the trial. One possible measure of this criterion for a trial with, say, r replicates is the mean squared error of a predicted difference [msepd($\mathbf{a}$, $\mathbf{m}$, r)], where

$$\text{msepd}(\mathbf{a}, \mathbf{m}, r) = E[(\mathbf{a}'\mathbf{y}_{i.} - \mathbf{a}'\mathbf{y}_{j.}) - (\mathbf{m}'\mathbf{v}_i - \mathbf{m}'\mathbf{v}_j)]^2. \tag{4}$$

Here E[.] denotes the expected value for two randomly chosen varieties i and j. It is easily shown that under the multivariate model (1) for the yields,

$$\text{msepd}(\mathbf{a}, \mathbf{m}, r) = 2/r\ \mathbf{a}'\Sigma_e\mathbf{a} + 2\ \mathbf{a}'\Sigma_v\mathbf{a} + 2\ \mathbf{m}'\Sigma_v\mathbf{m} - 4\ \mathbf{a}'\Sigma_v\mathbf{m}. \tag{5}$$

Thus when estimates of $\Sigma_v$ and $\Sigma_e$, as well as the coefficient vector $\mathbf{m}$ are available, several possible predictors ($\mathbf{a}$) and numbers of replicates (r) can be inserted into this equation to determine the msepd associated with various numbers of years and replicates.

To normalize the msepd values, they can be compared to the msepd value that would accrue from the most that could be expected of alfalfa breeders: a standard 4-replicate trial carried on for the full length of time specified by the long-term yield of interest (in other words, if $\mathbf{a}=\mathbf{m}$). We therefore define predictability (p) of differences among varieties for a specified long-term yield, using a given combination of short-term yields, as the efficiency (Cochran and Cox 1957) of the predicted yield differences based on short-term yield data, relative to that of a standard four replicate trial carried on for the full four, five or six years (as appropriate). Thus,

$$p = [\mathrm{msepd}(\mathbf{m}, \mathbf{m}, 4)/\mathrm{msepd}(\mathbf{a}, \mathbf{m}, r)]^{1/2} \times 100. \qquad (6)$$

Note that the patterns of variances and correlations can cause the value of p to exceed 100%.

## 2.5. Long-Term Yields of Interest

The predictability measure is general enough to be used to investigate whatever long-term yield is deemed to be of interest to alfalfa breeders. The choice of which long-term yields to investigate is an important issue, and involves economic as well as statistical principles. For this paper we will assume that the total yields over various ranges of years are the quantities to be used for selection of varieties. Other quantities to investigate in the future are the 'total discounted yields' over various ranges of years (Melton 1980).

Predictability in each of the trials was investigated for the total yields for years 1-4 and 1-5. In addition, for L85 and S85 the total yield for years 1-6 was investigated. Sets of coefficients (ie. coefficient vectors) associated with the long-term yields to be predicted ($\mathbf{m}$) and the short term yields to be used as predictors ($\mathbf{a}_i$) are summarized in Table 1. For each combination of long-term yields and number of production years used in the short-term predictors, two prediction vectors ($\mathbf{a}_1$ and $\mathbf{a}_2$) are given. The first of these is an equally weighted mean or total, and the second is weighted most heavily on the last of the prediction years. The weights for $\mathbf{a}_1$ and $\mathbf{a}_2$ in this paper were chosen arbitrarily. Other choices of weights should be considered in the future. Predictability (Eq. 6) was evaluated for 2 to 8 replicates in all cases.

# 3. Results

## 3.1. Variances and Covariances of Forage Yields

As examples, estimated parameters of $\Sigma_v$ and $\Sigma_e$ for L85 are given in Table 2. Estimated parameters for the other trials were similar. In all Lethbridge trials, the coefficients of variation (cv)of the random errors, calculated by dividing the diagonal elements of $\Sigma_e$ by the mean yearly yields, tended to be higher in later years of production than in the early years; however in S85, the cv increased for three years and then decreased for the final three years of production. In all four trials, correlation coefficients between errors for different years were usually highest for successive years, and decreased as the period between years increased. On average, the highest and lowest correlation coefficients between errors for any specified years were obtained in L84 and L85, respectively. In L83 and L84, correlations were generally higher among later production years than the early years while in L85 and S85, errors in production years 1-3 were generally more highly correlated with each other than with those in other years.

Treating varieties as random selections from some larger population, the cv among mean yields for varieties, calculated by dividing the diagonal elements of $\Sigma_v$ by the mean yearly yields, increased steadily with the years of production for L83 and L85, while in L84, the cv increased for four years and then decreased. No smooth trend for the cv was obtained in S85. In all four trials, the mean yields tended to be highly correlated between successive years, and generally decreased as the period between years increased.

### 3.2. Predictability

Predictabilities of the long-term total yields using specified linear combinations of observed short-term yields with various numbers of replicates are summarized in Table 3. This table gives two summary statistics for every combination of a long-term yield with a short-term predictor: 1) the predictability with four replicates (p) and 2) the number of replicates required to achieve 90% predictability (r). Predictability was arbitrarily considered to be good if $p \geq 90\%$ and/or $r \leq 4$.

Predictability of variety differences for the total four-year yield, using two different combinations of yields in the first three years of production, exceeded 100% with three or four replicates in both L83 and L84 (Table 3). For L85 and S85, using the equally-weighted combination of the three-year yields, predictability was 89% and 97%, respectively, with four replicates. For L84, L85, and S85, the equally weighted combination of the three-year yields was a more efficient predictor of the total four-year yield than the combination which gave more weight to the yield in the third production year.

With the exception of the unequally weighted predictor for L85, predictability of variety differences for the total five-year yield was lower than that for the total four-year yield in all trials (Table 3). Both combinations of yields in the first three years of production were highly efficient predictors of the total five-year yield in L83 and L84. In L84, the equally weighted combination of the three-year yields was over 100% efficient with four replicates, but the unequally weighted combination required five replicates to attain 90% efficiency. In L85 and S85, 90% predictability for the total five-year yield was achieved by using six and five replicates, respectively. In L85 and S85, predictability of variety differences for the total five-year yields increased markedly when yields from the first four years of production were used as predictors (Table 3). For both combinations of the four-year yields, at least 87% efficiency was attained with five replicates in both trials. For L83 and L84, predictability of the total five-year yield from the first four years was high, but was not much different from that obtained using the first three years of production.

Predictability of the total six-year yield using the first three years data, was generally poor in both L85 and S85, although 90% efficiency was achieved using one of the predictors with six replicates in S85 (Table 3). When the fourth year's data were added, predictability for L85 was still poor. In S85, however, predictability of the total six-year yield improved, and 90% efficiency was achieved by both predictors using five replicates. The addition of the fifth year's yields increased predictability in both L85 and S85. In L85, 90% efficiency was achieved using one of the predictors with six replicates while in S85, efficiency of both predictors was above 90% with only four replicates.

## 4. Discussion

Quantitative predictions of long-term yields of varieties in a trial, assessed by considering measures such as that of Eq. 6 of this paper, are needed to provide a quantitative basis for variety recommendations. Predictability of variety differences in their long-term yields varied from trial to trial, and was a complicated function of the matrices $\Sigma_e$ and $\Sigma_v$ for each trial. For example, the reason for the generally good predictability of variety differences for most long-term yields in L83 and L84 using only three years' data was different for each trial. In L83, the error cv was low in year three, whereas the error cv in years four and five was high. This, as well as the fact that variety means in years four and five were highly correlated with those in year three, made variety differences for long-term yields as predictable using short-term yields as if long-term data was available. In the case of L84, the high error cv in year three would be expected to lower the predictability of long-term variety differences. This effect was negated, however, by the fairly high correlation of plot errors among years. Predictability of variety differences for L85 was low because plot errors were not highly correlated among years, and the cv among variety means increased across production years. For S85, predictability was generally better than that for L85 because the opposite situation prevailed. Plot errors were correlated among years, and variability among variety means did not increase with age.

In spite of the differences among trials, some general statements can be made about predictability of alfalfa variety differences in their long-term yields. For total yields, predictors involving equally weighted combinations of short-term yields ($\mathbf{a}_1$ in Table 1) were in general more efficient than those involving combinations weighted most heavily on later yields ($\mathbf{a}_2$). Predictability of variety differences for the total 1-5 year yield was generally better if the predictor was based on four years' data rather than three. When four year's data are used, four replicates appear to be sufficient. On the other hand, variety differences for the total 1-6 year yield can only be adequately predicted if five years' data are available; even with five years' data, more than four replicates may be required to adequately predict cultivar differences. This is because in L85 at least 6 replicates were required to achieve 90% predictability.

This paper represents only a start in the process of accumulating information on how long alfalfa variety trials must be carried on, how many replicates are required, and what predictors should be used in predicting long-term variety differences. The statistical measures suggested in this paper need to be applied to more alfalfa yield trials for which long-term yield data are available. Nonetheless, based on the four trials available, it appears that variety differences for the total yield across four, five or six years of production can only be efficiently predicted if three, four, or five years' data, respectively, are available. Four replicates should generally be sufficient for predicting the total yield over four or five years of production, but more replicates may be needed for predicting the total six-year yield. This implies that current alfalfa forage yield trial designs should be either extended to five years and include six replicates, or extended to six years and include four replicates in order to obtain meaningful data for selecting varieties with high long-term (six-year) yields. If five years are used, predictions of the total six-year yield should be based on equally weighted combinations of yields in the five years of production rather than combinations weighted more heavily on the last year of production. It must be noted, however, that these recommendations should only be implemented after consideration of the costs associated with the various trial designs.

The methods developed in this paper might be extended or modified in various ways. For example: (1) These methods treat all comparisons of varietal means equally. However, variety trials usually include control varieties as well as experimental varieties, and comparisons of experimental to control varieties are of particular interest. An incorporation of this treatment structure into the current methodology may be profitable. (2) Comparisons of rankings of varieties may, in practice, be almost as useful as quantitative comparisons of their mean yields. Thus, it may be of interest to modify the current methods to consider only ranking of varieties. (3) If an acceptable value for msepd could be specified, Eq. 5 could be used to develop the relationship between r (the number of replicates) and the number of years represented in **a**. Such a relationship would be useful in studying the substitutability possible between replications and years. (4) Given a research budget and known costs per variety and year, it should be possible to find an 'optimal' combination of replications and years.

## 5. References

Cochran, W. G. and Cox, G. M. (1957) Experimental Designs. Wiley. New York.

Melton, B. E. (1980) Economics of beef cow culling and replacement decisions under genetic progress. West. J. Agric. Econ. 5:137.

Pearce, S. C. (1976) Field experimentation with fruit trees and other perennial crops, 2nd ed., Commonwealth Agricultural Bureaux, Slough, England.

SAS Institute (1989) SAS/STAT User's Guide, Version 6, Fourth Ed. Vol. 2, SAS Institute Inc., Cary, NC.

Stauber, S. and Goodman, G. (1986) Optimal replacement of alfalfa stands: a farm level decision model. Montana Ag. Research Spr/Sum 1986:1-5.

Stroup, W. W. and Mulitze, D. K. (1991) Nearest neighbor adjusted best linear unbiased prediction. Amer. Statist. 45:194-200.

**Table 1.** Coefficients of the total long-term yield vectors (**m**); and the equally ($\mathbf{a}_1$) and unequally ($\mathbf{a}_2$) weighted vectors used to predict long-term yields from short-term data.

| coef. of **m** for yr. | | | | | | prod. years used in | coef. of $\mathbf{a}_1$ for yr. | | | | | coef. of $\mathbf{a}_2$ for yr. | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | prediction | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 1 | 1 | 1 | 1 | 0 | 0 | 3 | 4/3 | 4/3 | 4/3 | 0 | 0 | 1 | 1 | 2 | 0 | 0 |
| 1 | 1 | 1 | 1 | 1 | 0 | 3 | 5/3 | 5/3 | 5/3 | 0 | 0 | 1 | 1 | 3 | 0 | 0 |
| | | | | | | 4 | 5/4 | 5/4 | 5/4 | 5/4 | 0 | 1 | 1 | 1 | 2 | 0 |
| 1 | 1 | 1 | 1 | 1 | 1 | 3 | 2 | 2 | 2 | 0 | 0 | 1 | 1 | 4 | 0 | 0 |
| | | | | | | 4 | 3/2 | 3/2 | 3/2 | 3/2 | 0 | 1 | 1 | 1 | 3 | 0 |
| | | | | | | 5 | 6/5 | 6/5 | 6/5 | 6/5 | 6/5 | 1 | 1 | 1 | 1 | 2 |

**Table 2.** Coefficients of variation (diagonal) and correlation coefficients (off-diagonal) associated with the error covariance matrix ($\Sigma_e$) and the variety covariance matrix ($\Sigma_v$) for the 6-year series of yields from the L85 trial.

| matrix | prod. year | production year 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $\Sigma_e$ | 1 | 4.7 | 0.49 | 0.59 | 0.36 | 0.10 | -0.18 |
| | 2 | | 6.5 | 0.49 | 0.27 | 0.07 | 0.07 |
| | 3 | | | 10.4 | 0.57 | 0.18 | -0.24 |
| | 4 | | | | 7.5 | 0.48 | 0.20 |
| | 5 | | | | | 14.6 | 0.53 |
| | 6 | | | | | | 13.8 |
| $\Sigma_v$ | 1 | 4.7 | 0.85 | 0.96 | 0.91 | 0.86 | 0.91 |
| | 2 | | 4.8 | 1.00 | 0.87 | 0.67 | 0.71 |
| | 3 | | | 3.9 | 1.00 | 0.85 | 1.00 |
| | 4 | | | | 7.5 | 0.94 | 0.92 |
| | 5 | | | | | 13.4 | 1.00 |
| | 6 | | | | | | 21.2 |

**Table 3.** Predictability (p) of cultivar differences for total forage yield over production years 1-4, 1-5 and 1-6 using two short-term predictors[1] with 4 replications; and number of replicates (r) needed to achieve 90% predictability, for the alfalfa trials.

| | | L83[2] | | | | L84 | | | | L85 | | | | S85 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **a**$_1$ | | **a**$_2$ | | **a**$_1$ | | **a**$_2$ | | **a**$_1$ | | **a**$_2$ | | **a**$_1$ | | **a**$_2$ | |
| years of predicted total | num. years used for prediction | p | r | p | r | p | r | p | r | p | r | p | r | p | r | p | r |
| 1-4 | 3 | 113 | 3 | 118 | 3 | 116 | 3 | 100 | 4 | 89 | 5 | 76 | 6 | 97 | 4 | 87 | 5 |
| 1-5 | 3 | 98 | 3 | 110 | 2 | 108 | 3 | 87 | 5 | 75 | 6 | 61 | *[3] | 84 | 5 | 70 | 8 |
| | 4 | 111 | 3 | 103 | 3 | 93 | 4 | 87 | 5 | 88 | 5 | 93 | 4 | 90 | 4 | 87 | 5 |
| 1-6 | 3 | | | | | | | | | 56 | * | 49 | * | 79 | 6 | 63 | * |
| | 4 | | | | | | | | | 63 | * | 70 | * | 85 | 5 | 81 | 5 |
| | 5 | | | | | | | | | 76 | 8 | 81 | 6 | 96 | 4 | 101 | 4 |

[1] **a**$_1$ was an equally weighted linear combination of the yields in all production years used for prediction; **a**$_2$ was a linear combination of the yields with greatest weight on the last of the production years used for prediction. See Table 1 for details.

[2] L83, L84, and L85 were grown in Lethbridge, Alberta; S85 was grown in Swift Current, Saskatchewan.

[3] did not achieve 90% predictability with 8 replications.