THE INSTITUTE OF STATISTICS

UNIVERSITY OF NORTH CAROLINA SYSTEM

REPEATED MEASURES IN RANDOMIZED BLOCK AND SPLIT PLOT EXPERIMENTS

M. L. Gumpertz and C. Brownie

Institute of Statistics Mimeograph Series No. 2202

November, 1991

NORTH CAROLINA STATE UNIVERSITY
Raleigh, North Carolina

Repeated Measures in Randomized Block and Split Plot Experiments

M. L. Gumpertz and C. Brownie

Department of Statistics
North Carolina State University
Box 8203
Raleigh, NC 27695-8203
(919)515-2584

Gumpertz, M. L. and Brownie, C. 1991. Repeated measures in randomized block and split plot experiments.

Randomized block and split plot designs are among the most commonly used experimental designs in forest research. Measurements for plots in a block (or subplots in a whole plot) are correlated with each other, and these correlations must be taken into account when analyzing repeated measures data from blocked designs. The analysis is similar to repeated measures analysis for a completely randomized design, but test statistics must allow for random block × time effects, and standard errors for treatment means must also incorporate block-to-block variation and block × treatment interactions as well as variation among plots within a block. Two types of statistical analysis are often recommended for repeated measures data: analysis of contrasts of the repeated factor, and multivariate analysis of variance. A complete analysis of repeated measures should usually contain both of these components, just as in univariate analysis of variance it is often necessary to decompose the main effects into single degree of freedom contrasts to answer the research objectives. We demonstrate the multivariate analysis of variance and the analysis of contrasts in detail for two experiments. In addition, estimation of coefficients assuming a polynomial growth curve is discussed in detail for one of these experiments. The first experiment, a randomized complete block design, is a forest nutrition study of the long term effects of mid-rotation nitrogen and phosphorus fertilization on loblolly pine (Pinus taeda L.), and the second, a split plot design, is an air pollution study of the effects of ozone and acid precipitation on loblolly pine growth.

Introduction

Plant scientists are becoming increasingly aware of the need for careful statistical analysis when the data collected represent a series of measurements over time or space on each experimental unit. Evidence of this growing interest in the analysis of repeated measures data is provided by a number of recent articles, e.g., Johnson, Chaudhuri, and Kanemasu (1983), Eskridge and Stevens (1987), Moser, Saxton and Pezeshki (1990), and Meredith and Stehman (1991). The papers by Moser et al. (1990) and Meredith and Stehman (1991) focus specifically on repeated measures data in forestry and give good examples of experiments which yield this type of data. For the situation where treatments are assigned to experimental units according to a completely randomized design (CRD), Moser et al. describe a model and demonstrate multivariate analysis of variance (MANOVA), while Meredith and Stehman recommend "analysis of coefficients".

Two issues do not seem to have been adequately addressed in any of these articles. First, the appropriate models and methods of analysis have not been explicitly described for experiments where some factors are random, such as randomized block designs where blocks are considered to be random replication. This seems a serious omission since a substantial proportion of field and greenhouse trials in forestry and agriculture utilize a blocked design such as a randomized complete block design (RCBD) or a split plot design. Second, there is some confusion concerning the labeling of analyses as either "analysis of coefficients" (equivalently, analysis of contrasts) or as MANOVA. Accordingly, in this article we have two objectives. We describe appropriate models and analyses for repeated measures data for randomized complete block and split plot designs. Also we try to reconcile the different recommendations concerning analysis of contrasts and MANOVA by showing that the former can be thought of as part of a MANOVA approach.

Randomized complete block design with repeated measures

Experimental situation and data

We focus on experiments laid out according to a randomized complete block design with b

blocks and *t* treatments, where in each of the *bt* experimental units, a measurement is recorded at each of *p* points in time, or at each of *p* locations. The times or locations of the *p* repeated measures must be consistent across the experimental units, and we refer to the times or locations as levels of the repeated measures factor. As in Meredith and Stehman (1991), we limit consideration to studies where the repeated measures factor is a systematic factor such as time, distance, or soil depth, with levels that cannot be assigned at random. For convenience we will often refer to the repeated measures factor as "time".

The systematic nature of the repeated measures factor has important consequences for data analysis and interpretation of results. One consequence relates to the effect on the correlation structure of the data and appropriate statistical tests. A second consequence (which is often overlooked) is the potential for confounding between effects of the repeated measures factor and concomitant environmental effects over which the researcher has no control. For example, if biweekly measurements are taken on plant height and there is a sudden warm wet spell, age and environmental effects will both contribute to the shape of the growth curve and there may be no way to separate the two. Similarly, distance from an irrigation source could be confounded with underlying patterns in the soil variation. The possibility that effects attributed to a repeated measures factor could be due in part to unknown environmental variation should always be kept in mind when analyzing repeated measures data.

Johnson et al. (1983) give an example where the repeated measures factor is distance from an irrigation source. We consider an example where the repeated measures factor is time. This example is taken from a study of the effects of mid-rotation nitrogen (N) and phosphorus (P) fertilization on loblolly pine (Pinus taeda L.) stands at several sites across the southeastern United States that was carried out by the North Carolina State Forest Nutrition Cooperative (Valentine and Allen 1990). We selected one site to illustrate repeated measures analysis for the effects of N and P on tree volume over time. Plots were arranged in a randomized block design with 12 plots per block. The twelve treatments were factorial combinations of four levels of N and three levels of

P. Diameter and height of all trees in the plot were measured at two, four, and six years after the fertilizer application to estimate tree volume per acre. Thus the repeated measures factor, time, has p=3 levels (2, 4, and 6 years post-application), and for each plot there is an estimate of tree volume at each of the 3 times (see Table 1).

Model, assumptions, and notation

Rowell and Walters (1976) and Moser et al. (1990), have explained why the systematic nature of the repeated measures factor invalidates the traditional univariate analyses for repeated measures data. We extend these discussions to include consideration of the correlation structure of the data when random block effects are present.

The traditional univariate analysis for repeated measures data treats the repeated measurements as though they were subplots in a split plot design. It assumes that observations within a whole plot unit are equicorrelated. In a true split plot design it might seem that the correlation between two observations should depend on the distance between the subplots, but the random assignment of treatments to subplots provides justification for the assumption of equal correlations (e.g., Kempthorne, 1952, Chapters 7-9). Briefly, this can be seen by calculating the correlation between values for any two treatments within a whole plot, under the null hypothesis of no treatment effects, and with respect to all equally likely assignments of treatments to subplots within a plot. For repeated measures data, it is not possible to appeal to a randomization argument to justify the "equal correlations" assumption for observations on the same plot, because levels of the within-plot factor (location or time) cannot be assigned randomly. In fact the systematic nature of the repeated measures factor often results in stronger correlations between adjacent observations than between observations that are well separated, either in space or time. As noted by Moser et al., this failure of the "equal correlations" assumption for observations within a plot invalidates a split plot or split block analysis.

When plots are arranged in blocks, systematic allocation of levels of the repeated measures

factor can also induce correlations between the random block × time (or block × location) interaction effects. That is, the effect of the i^{th} block at a particular time is correlated with its effect at any other time. Looked at another way, the effect is modified by each block and the magnitudes of these random block × time "errors" are correlated across measurement times. Thus, instead of the more common assumption of independent block × time effects, we must allow for (possibly different) correlations between these interaction effects within the same block. Having allowed for correlations to differ between pairs of times, it is natural to allow variances to change across time also. This is particularly realistic in studies where plants are measured at different stages of growth.

The resulting multivariate repeated measures model is

$$[1] \quad y_{ijk} = \underbrace{\mu + \beta_i + \mathrm{A}_j + \epsilon_{ij}}_{\text{between-plot}} + \underbrace{\mathrm{T}_k + (\mathrm{AT})_{jk} + (\beta\mathrm{T})_{ik} + \delta_{ijk}}_{\text{within-plot}},$$

with the usual restrictions: $\sum_{j=1}^{a} \mathrm{A}_j = \sum_{k=1}^{p} \mathrm{T}_k = \sum_{j=1}^{a}\sum_{k=1}^{p} (\mathrm{AT})_{jk} = \sum_{i=1}^{p} (\beta\mathrm{T})_{ik} = 0$. **The "between-plot" or "between-subject" part of the model consists of an overall mean, μ, a random block effect, β_i, a fixed treatment effect, A_j, and random plot to plot variation, ϵ_{ij}. This part of the model is just the traditional model for a randomized block experiment. The block effects are assumed to have variance σ^2_β and to be uncorrelated with each other and the plot errors are uncorrelated from plot to plot with variance σ^2_ϵ. These assumptions imply that observations on two plots in the same block are correlated but that two plots in separate blocks are uncorrelated with each other.**

The "time", "repeated measures", or "within-plot" part of the model consists of T_k, a time effect, $(\mathrm{AT})_{jk}$, a time × treatment interaction, $(\beta\mathrm{T})_{ik}$, a random interaction of block and time, and δ_{ijk}, a random effect for observations on the same plot. In this repeated measures model we assume that the within-plot effects, δ_{ijk}, are correlated, with variance at time k denoted $\mathrm{Var}(\delta_{ijk}) = \sigma_{\delta kk}$ and covariance between time k and time k^* denoted $\mathrm{Cov}(\delta_{ijk}, \delta_{ijk^*}) = \sigma_{\delta kk^*}$.

Within each block the block × time interactions are also correlated and the correlation may be different between times 1 and 2, for instance, than between times 1 and 3. This is denoted by $\mathrm{Cov}(\beta T_{ik}, \beta T_{ik^*}) = \sigma_{\beta T k k^*}$. For purposes of testing hypotheses and constructing confidence intervals, all random effects are assumed to be normally distributed.

If the p responses for a single plot are arranged in a row vector, $\underset{\sim}{y}_{ij}$, we can write the repeated measures model in a form which is analogous to the familiar randomized block analysis of variance model. In vector notation the repeated measures model takes the following simple and compact form, in which each term represents a vector of p elements corresponding to the p repeated measures:

[2] $\underset{\sim}{y}_{ij} = \underset{\sim}{\mu} + \underset{\sim}{\beta}_i + \underset{\sim}{A}_j + \underset{\sim}{\epsilon}_{ij}$,

where i=1,...,b blocks and j=1,...,a levels of factor A. The correspondence between the vector notation in expression [2] and the notation in [1] is as follows:

$$\underset{\sim}{y}_{ij} = [y_{ij1} \;\; y_{ij2} \;\; \cdots \;\; y_{ijp}],$$
$$\underset{\sim}{\mu} = [\mu + T_1 \;\; \mu + T_2 \;\; \cdots \;\; \mu + T_p],$$
$$\underset{\sim}{\beta}_i = [\beta_i + (\beta T)_{i1} \;\; \beta_i + (\beta T)_{i2} \;\; \cdots \;\; \beta_i + (\beta T)_{ip}],$$
$$\underset{\sim}{A}_j = [A_j + (AT)_{j1} \;\; A_j + (AT)_{j2} \cdots A_j + (AT)_{jp}],$$
$$\underset{\sim}{\epsilon}_{ij} = [\epsilon_{ij} + \delta_{ij1} \;\; \epsilon_{ij} + \delta_{ij2} \;\; \cdots \;\; \epsilon_{ij} + \delta_{ijp}].$$

Here the block effects $\underset{\sim}{\beta}_i$ are assumed to have covariance $\underset{\sim}{\Sigma}_\beta$ and the plot-to-plot variations, $\underset{\sim}{\epsilon}_{ij}$, have covariance $\underset{\sim}{\Sigma}_\epsilon$. The two covariance matrices are completely unstructured; that is, the variances are allowed to be different for different times and covariances between different times can be different for all pairs of times. By relating terms in the models [2] and [1] we see that the matrix $\underset{\sim}{\Sigma}_\beta$ contains components due to the covariance between plots in a block and to block × time interactions; the kk^{*th} element is $\sigma^2_\beta + \sigma_{\beta T k k^*}$. The matrix $\underset{\sim}{\Sigma}_\epsilon$ contains variances and covariances among observations on one plot , and its kk^{*th} element is $\sigma^2_\epsilon + \sigma_{\delta k k^*}$. The covariances among the p repeated measures are:

observations on the same plot: $\underset{\sim}{\mathrm{Var}}(\underset{\sim}{y}_{ij})=\underset{\sim}{\Sigma}_{\beta}+\underset{\sim}{\Sigma}_{\epsilon}$,

two different plots in the same block: $\underset{\sim}{\mathrm{Cov}}(\underset{\sim}{y}_{ij},\underset{\sim}{y}_{ij^*}) = \underset{\sim}{\Sigma}_{\beta}$,

two plots in different blocks: $\underset{\sim}{\mathrm{Cov}}(\underset{\sim}{y}_{ij},\underset{\sim}{y}_{i^*j^*}) = \underset{\sim}{0}$.

As noted above, the covariance structure assumes neither constant variance across time nor equicorrelated errors between pairs of times. However the model does contain a strong homogeneity of variance assumption in that $\underset{\sim}{\Sigma}_{\epsilon}$ and $\underset{\sim}{\Sigma}_{\beta}$ are assumed constant across plots. In other words, under model [2], the covariance heterogeneity is associated entirely with the levels of the repeated measures factor, and consequently the variance assumptions for the usual RCBD analysis hold for the data from a single measurement time. Similarly, the RCBD variance assumptions hold for any linear combination of the repeated measurements (such as the plot mean $\overline{y}_{ij.} = \frac{1}{p} \sum_{k=1}^{p} y_{ijk}$) computed for each plot. The analysis of contrasts described in the next section utilizes this aspect of the covariance structure of the model [2].

Analysis Methods

For data like those in Table 1, a common approach is to do a separate analysis of variance for each of the three measurement dates. These individual analyses yield valid tests for treatment (i.e., fertilizer) effects at a fixed point in time but (see e.g., Rowell and Walters, 1976; Meredith and Stehman, 1991) this approach is unsatisfactory because it does not yield formal tests for assessing trends over time or time × treatment interactions. Rowell and Walters (1976) instead proposed the "analysis of contrasts" for examining time and time × treatment interaction effects. More recently, Moser et al. (1990) demonstrated multivariate analysis of variance for repeated measures data from a completely randomized design. For the randomized complete block design we first describe the analysis of contrasts and then show how this can be carried out in the context of a multivariate analysis of variance for repeated measures data.

Analysis of contrasts

The (repeated measures) analysis of contrasts can be thought of as a series of analyses of variance, each on a different linear combination of the measurements across time. First, an analysis of variance is applied to the plot means $\overline{y}_{ij.}$. This is called the "between-plot" analysis and provides a test of treatment effects averaged over all times. The second component of the analysis, the "within-plot" analysis, consists of separate analyses on each of p-1 contrasts (linear combinations with coefficients that sum to 0) on the repeated measurements. The p-1 contrasts should reflect questions of interest concerning the time effect. If the relationship between the response and time can be approximated with a simple polynomial, then orthogonal polynomial contrasts would be selected, as in the analysis of coefficients described by Meredith and Stehman (1991). The linear contrast is evaluated from the p repeated measurements in each plot, and an analysis of variance is applied to these values, with a term for the mean included. Similarly, analyses of variance are carried out on the quadratic and higher order polynomial contrasts. The analysis of variance on the linear contrast provides tests of whether the trend across time has a linear component (averaging over treatments) and of whether the linear component is the same for each treatment, and analogously for the quadratic and higher order polynomial contrasts.

Results for the between- and within-plot analyses are presented in Table 2 for the fertilizer example (SAS® code is in Appendix I). The fertilizer effects are partitioned into N, P and N × P effects reflecting the factorial treatment structure. The between-plot analysis indicates that, averaged over 2, 4, and 6 years post-application there is a strong main effect of N on tree volume.

Trends over the 6 year period are examined in the within-plot analysis. As there are three equally spaced measurement times, values of the linear contrast are obtained on each plot as $z_{ij1} = \frac{1}{\sqrt{2}}(y_{ij1} - y_{ij3})$, where the multiplier $\frac{1}{\sqrt{2}}$ simply ensures that sums of squares for the linear and quadratic contrasts will represent a partitioning of the overall sum of squares for time. The quadratic contrast is $z_{ij2} = \frac{1}{\sqrt{6}}(y_{ij1} - 2y_{ij2} + y_{ij3})$. Results for the analysis of variance of the z_{ij1} values are listed first, and provide information on the role of the linear component in a regression of tree volume on time. The analysis of variance of the quadratic contrast values, z_{ij2}, indicates whether the response over time is curved rather than linear, and whether this nonlinear component is affected by the fertilizer treatment.

Each fixed effect is tested against its interaction with blocks so that conclusions will not be confined to the blocks in this particular experiment. The proper error term for testing fertilizer differences is based on the variability in fertilizer effects among blocks. Similarly, the proper error terms for testing the time contrasts are based on the variability in those contrasts from block to block, i.e. the contrast × block interaction. Interpretation of these F tests is discussed following the results of the MANOVA tests.

Testing hypotheses: multivariate analysis

The analysis of contrasts can be viewed as a univariate analysis in which contrasts among times are used to partition sums of squares for time and time × fertilizer effects, and the corresponding error terms, in order to construct valid F tests. Alternatively the analysis of contrasts can be viewed as a part of a multivariate repeated measures analysis. The multivariate tests for the fertilizer effects (averaged over time) are the same as the tests in the between-plot part of the analysis of contrasts (Table 2). In addition, the single degree of freedom time contrasts are combined into composite multivariate tests for the time main effect and time × fertilizer interactions. These multivariate tests for the time main effect and the time × fertilizer interactions jointly examine hypotheses relating to all of the p-1 contrasts on time. In the fertilizer example, the test for the time main effect puts the linear and quadratic time effects together into one hypothesis.

In order to see exactly what hypotheses are usually tested in a repeated measures analysis we put all of the parameters of the model into one matrix, $\underset{\sim}{\mathbf{B}}$, with 17 rows (one for the mean, four for block effects and 12 for fertilizer effects) and 3 columns, one for each time.

$$\underset{\sim}{\mathbf{B}} = \begin{bmatrix} \mu + T_1 & \mu + T_2 & \mu + T_3 \\ \beta_1 + (\beta T)_{11} & \beta_1 + (\beta T)_{12} & \beta_1 + (\beta T)_{13} \\ \vdots & \vdots & \vdots \\ \beta_4 + (\beta T)_{41} & \beta_4 + (\beta T)_{42} & \beta_4 + (\beta T)_{43} \\ F_1 + (FT)_{1,1} & F_1 + (FT)_{1,2} & F_1 + (FT)_{1,3} \\ \vdots & \vdots & \vdots \\ F_{12} + (FT)_{12,1} & F_{12} + (FT)_{12,2} & F_{12} + (FT)_{12,3} \end{bmatrix} \begin{matrix} \text{mean} \\ \text{block 1} \\ \vdots \\ \text{block 4} \\ \text{fertilizer 1} \\ \vdots \\ \text{fertilizer 12} \end{matrix} .$$

The hypotheses are then all written in the form H_0: $\underset{\sim}{L}\,\underset{\sim}{B}\,\underset{\sim}{M} = 0$, where the matrix $\underset{\sim}{L}$ specifies the linear combinations of the between-plot factors and the matrix $\underset{\sim}{M}$ specifies the linear combinations of the within-plot factors to be tested. Let $\underset{\sim}{L}_n$ contain coefficients for comparing nitrogen levels, $\underset{\sim}{L}_p$ contain coefficients for comparing phosphorus levels, $\underset{\sim}{L}_{np}$ contain coefficients for nitrogen × phosphorus interaction, and $\underset{\sim}{M}_t$ contain time contrast coefficients. In addition, let $\underset{\sim}{M}_0$ contain coefficients to compute the mean of all times and $\underset{\sim}{L}_0$ coefficients to compute the mean over all plots (for examples of $\underset{\sim}{L}$ and $\underset{\sim}{M}$ matrices see Appendix II). These contrast matrices are put together to test the repeated measures hypotheses as follows :

Nitrogen main effect	H: $\underset{\sim}{L}_n \underset{\sim}{B}\, \underset{\sim}{M}_0 = \underset{\sim}{0}$	Compare N levels averaged over all levels of P and time
Phosphorus main effect	H: $\underset{\sim}{L}_p \underset{\sim}{B}\, \underset{\sim}{M}_0 = \underset{\sim}{0}$	Compare P levels averaged over all levels of N and time
N × P interaction	H: $\underset{\sim}{L}_{np} \underset{\sim}{B}\, \underset{\sim}{M}_0 = \underset{\sim}{0}$	N × P interaction averaged over all times
Time main effect	H: $\underset{\sim}{L}_0 \underset{\sim}{B}\, \underset{\sim}{M}_t = \underset{\sim}{0}$	Compare times averaged over fertilizer levels
Time × Nitrogen	H: $\underset{\sim}{L}_n \underset{\sim}{B}\, \underset{\sim}{M}_t = \underset{\sim}{0}$	Time × N interaction (averaged over levels of P)
Time × Phosphorus	H: $\underset{\sim}{L}_p \underset{\sim}{B}\, \underset{\sim}{M}_t = \underset{\sim}{0}$	Time × P interaction (averaged over levels of N)
Time × N × P	H: $\underset{\sim}{L}_{np} \underset{\sim}{B}\, \underset{\sim}{M}_t = \underset{\sim}{0}$	Time × N × P interaction

These hypotheses have the same meanings as analysis of variance hypotheses. In particular, a main effect is averaged over the levels of all the other factors. Note that these hypotheses may be different from the hypotheses tested in a general purpose MANOVA program (as opposed to a program that is specifically intended for repeated measures analysis). In a repeated measures analysis the tests for fertilizer effects compare levels of N and/or P averaged over all times, whereas in a general MANOVA the tests for fertilizer effects compare fertilizer levels for all times simultaneously. The test for the time main effect uses the contrasts of time, z_{ij1} and z_{ij2}, jointly, to test that the mean response over all plots is the same at each time. Any pair of linearly independent contrasts over time would result in the same multivariate tests for time main effect and interactions as the linear and quadratic contrasts presented here.

In the univariate analysis of variance for a randomized block design, hypotheses are tested by comparing the treatment mean squares with MSBlock × Treatment, which estimates the residual variance or error. In the repeated measures setting we have $p \times p$ matrices of treatment and error mean squares. The multivariate tests are based upon these matrices of sums of squares and cross products, which play the same role as the sums of squares in a univariate analysis. For the randomized complete block design with one "between-plot" factor, A, let $\mathrm{SSB \times A}_{y_k}$ be the block × A sum of squares for the k^{th} time. $\mathrm{SSB \times A}_{y_k}$ is based on the residuals $r_{ijk} = y_{ijk} - \bar{y}_{i.k} - \bar{y}_{.jk} + \bar{y}_{..k}$. The sum of products between the residuals at the k^{th} and k^{*th} times is denoted $\mathrm{SPB \times A}_{y_k y_{k^*}} = \sum_{i=1}^{4} \sum_{j=1}^{12} r_{ijk}\, r_{ijk^*}$. These sums of squares and cross products are arranged in the matrix

$$\underset{\sim}{\mathrm{SSCP}}(\mathrm{Block \times A})_y = \begin{bmatrix} \mathrm{SSB \times A}_{y_1} & \mathrm{SPB \times A}_{y_1 y_2} & \mathrm{SPB \times A}_{y_1 y_3} \\ \mathrm{SPB \times A}_{y_2 y_1} & \mathrm{SSB \times A}_{y_2} & \mathrm{SPB \times A}_{y_2 y_3} \\ \mathrm{SPB \times A}_{y_3 y_1} & \mathrm{SPB \times A}_{y_3 y_2} & \mathrm{SSB \times A}_{y_3} \end{bmatrix}.$$

Sum of squares and cross product matrices are constructed for each of the terms in the model [2] in a similar fashion. For the randomized block design with one between-plot factor there are four such matrices (Table 3).

The within-plot covariance matrix, $\underset{\sim}{\Sigma}_{\epsilon}$, is estimated by the Block × A SSCP matrix divided by its degrees of freedom. The matrix $\frac{1}{b\text{-}1}\, \underset{\sim}{\mathrm{SSCP}}(\mathrm{Block})_{\mathrm{y}}$ estimates a linear combination of the block and within-plot covariance matrices, $a \cdot \underset{\sim}{\Sigma}_{\beta} + \underset{\sim}{\Sigma}_{\epsilon}$. This is a direct generalization of expected mean squares in the univariate analysis of variance with random block effects, and the univariate expected mean squares can be used as a guide to determine the proper error matrix for a particular test.

The multivariate tests can be constructed directly from the matrices in Table 3 (see e.g. Johnson et. al. 1983). However, the relationship between the analysis of contrasts and the multivariate tests is more readily seen if we use the $p-1$ contrast variables, z_{ijl}, ℓ=1,...p-1, and the normalized mean, $z_{ij0} = \frac{1}{\sqrt{p}} \sum_{k=1}^{p} y_{ijk}$, rather than the original variables, $y_{ij1}, ... y_{ijp}$, to construct the matrices of sums of squares and products. Notice that the contrast variables can be formed by post-multiplying $\underset{\sim}{y}$ by a contrast matrix $\underset{\sim}{M}$: $z_{ij0} = \underset{\sim}{y}_{ij} \underset{\sim}{M}_0$ and $[z_{ij1} \; z_{ij2}] = \underset{\sim}{y}_{ij} \underset{\sim}{M}_t$. The tests for between-plot effects are based on the sums of squares of z_{ij0}. The multivariate tests of time main effects and interactions are based on the sums of squares of the p - 1 contrasts, z_{ij1}, z_{ij2}, ... , z_{ijp-1}, defined in the matrix $\underset{\sim}{M}_t$. In the fertilizer experiment z_{ij1} is the linear contrast, and z_{ij2} is the quadratic contrast. For the ℓ^{th} contrast denote the sum of squares for a particular effect SSEffect$_{zl}$, and denote the sum of cross products between the first and second contrasts SPEffect$_{z1z2}$. These sums of squares and cross products are arranged in the matrix

$$\underset{\sim}{\text{SSCP}}(\text{Effect})_z = \begin{bmatrix} \text{SSEffect}_{z1} & \text{SPEffect}_{z1z2} \\ \text{SPEffect}_{z1z2} & \text{SSEffect}_{z2} \end{bmatrix}.$$

The test for the time main effect compares the determinant of $\underset{\sim}{\text{SSCP}}(\text{Mean})_z$, which contains sums of squares and cross products of the grand means $\bar{z}_{..1}$ and $\bar{z}_{..2}$ of the linear and quadratic time contrast variables, with the determinant of $\underset{\sim}{\text{SSCP}}(\text{Block})_z$, which contains sums of squares for block-to-block variation. The determinant, which is also called the generalized variance, summarizes the information about the variances of both the linear contrast and the quadratic contrast and their covariance. The likelihood ratio test statistic is Wilks' lambda,

$$\Lambda = \frac{|\underset{\sim}{\text{SSCP}}(\text{Block})_z|}{|\underset{\sim}{\text{SSCP}}(\text{Mean})_z + \underset{\sim}{\text{SSCP}}(\text{Block})_z|}.$$

The hypothesis of no time main effect is rejected if Λ is small; i.e., if the block-to-block variation in time effects is small relative to mean differences in tree volume among times.

In general, Wilks' lambda compares an error matrix ($\underset{\sim}{E}$) to a hypothesis matrix ($\underset{\sim}{H}$), which indicates which of the "between-plot" factors are being tested. In a univariate analysis of variance for a randomized block design, the fixed treatment factors are tested against the mean square for interaction with blocks. Similarly, in the repeated measures case, the error matrix for testing a "between-plot" factor is the block × "between-plot" factor $\underset{\sim}{SSCP}$ matrix. The general form of Wilks' lambda is

$$\Lambda = \frac{|\underset{\sim}{E}|}{|\underset{\sim}{H} + \underset{\sim}{E}|}$$

For the hypothesis that the time main effect equals zero, $\underset{\sim}{H} = \underset{\sim}{SSCP}(\text{Mean})_z$ and $\underset{\sim}{E} = \underset{\sim}{SSCP}(\text{Block})_z$. The matrices for testing fertilizer × time interactions are $\underset{\sim}{H} = \underset{\sim}{SSCP}(\text{Fertilizer})_z$ and $\underset{\sim}{E} = \underset{\sim}{SSCP}(\text{Block} \times \text{Fertilizer})_z$. The fertilizer main effects could also be tested using a Wilks' lambda statistic with $\underset{\sim}{H} = \underset{\sim}{SSCP}(\text{Fertilizer})_{z_0}$ and $\underset{\sim}{E} = \underset{\sim}{SSCP}(\text{Block} \times \text{Fertilizer})_{z_0}$, but since only one variable, z_0, is involved, the matrices $\underset{\sim}{E}$ and $\underset{\sim}{H}$ are scalars (each involving only a single sum of squares) and the test statistic reduces to the usual F statistic, $\text{MSFertilizer}_{z_0}/\text{MSBlock} \times \text{Fertilizer}_{z_0}$. The distribution of Wilks' lambda depends on the hypothesis degrees of freedom, d_h, the error degrees of freedom, d_e, and on the number of contrasts for the within-plot factor, $c=p-1$. The test can only be carried out if $d_e+1 > c$. In a randomized complete block design with time as the repeated factor, the number of blocks must exceed the number of time contrasts. Critical values for hypothesis tests are given in Appendix III.

The matrices and the statistics for testing the time main effect and the fertilizer × time interactions in the fertilizer experiment follow. Compare the diagonal elements of the SSCP matrices with the sums of squares for the linear and quadratic contrasts of Table 2.

Time main effect:

$$\underset{\sim}{\mathrm{SSCP}}(\mathrm{mean})_y = \begin{bmatrix} 42264950 & 563133 \\ 563133 & 7503 \end{bmatrix} \qquad \underset{\sim}{\mathrm{SSCP}}(\mathrm{block})_y = \begin{bmatrix} 110624 & 9353 \\ 9353 & 3305 \end{bmatrix}$$

Λ=.002135, p-1=2, d_h=1, d_e=3

test statistic=467.3, p=.002

Fertilizer × Time interactions:

$$\underset{\sim}{\mathrm{SSCP}}(\mathrm{Block}\times\mathrm{Fertilizer})_z = \begin{bmatrix} 182685 & 5819 \\ 5819 & 28115 \end{bmatrix} \qquad p\text{-}1=2,\ d_e=4$$

$$\underset{\sim}{\mathrm{SSCP}}(\mathrm{N})_z = \begin{bmatrix} 461468 & -6101 \\ -6101 & 2720 \end{bmatrix} \quad \underset{\sim}{\mathrm{SSCP}}(\mathrm{P})_z = \begin{bmatrix} 47871 & 14453 \\ 14453 & 6088 \end{bmatrix} \quad \underset{\sim}{\mathrm{SSCP}}(\mathrm{N}\times\mathrm{P})_z = \begin{bmatrix} 24029 & 8499 \\ 8499 & 4500 \end{bmatrix}$$

N	P	N × P
d_h=3, Λ=.2569	d_h=2, Λ=.6826	d_h=6, Λ=.7805
test statistic=10.38	test statistic=3.366	test statistic=.7034
p=.0001	p=.015	p=.74

The linear and quadratic effects of time (averaged over fertilizer levels) from Table 2 are tested simultaneously by the Wilks' Λ test for time main effects. There are significant changes in tree volume over time (Wilks' lambda for time main effect, p=.002), and the interaction tests indicate that the change over time depends on both the N and P fertilization rates (Wilks' lambda, p=.0001 for N × time interaction and p=.015 for P × time interaction). From the analysis of contrasts in Table 2 we see that the rate of increase of tree volume depends on N (F-test for N × time linear effect, p=.0001). It also appears that the increase in volume with time is not linear and the amount of curvature depends on the level of P (F-test for P × time quadratic effect, p=.039). In the original analysis (Vose and Allen 1988), increase in volume from the time of fertilizer application was analyzed rather than volume itself. The multivariate tests for time and time × fertilizer effects and the tests of time contrasts are unaffected by subtracting initial volume from every measurement because they are based on differences among times. In other analyses

(Valentine and Allen 1990) the authors have used volume at the time of fertilization as a covariate. Volume at time zero can be incorporated into the repeated measures analysis as a "between-plot" covariate. In this particular data set inclusion of the covariate has the effect of substantially decreasing both the unexplained variation among blocks and the residual error and of drawing out a significant P main effect. We omit the covariate here simply to demonstrate the repeated measures method of analysis. In this data set the variance increases with time, which usually is an indication that a log transformation should be used to stabilize the variances; however in the repeated measures analysis there is no requirement that variances be constant across time, so the analysis can be performed in the original scale, making interpretation and presentation of results simpler.

Estimation of contrasts or polynomial coefficients

In addition to testing hypotheses, estimation of the treatment means of the variables $z_{ij0}, \ldots, z_{ijp-1}$ is also important (see also Meredith and Stehman 1991). In our example the z variables are orthogonal polynomial contrast variables, so the mean of z_{ij1} for the j^{th} fertilizer treatment ($\bar{z}_{.j1}$) gives the linear effect for the j^{th} fertilizer level, while the mean for z_{ij2} ($\bar{z}_{.j2}$) gives the estimated quadratic effect. The fact that blocks are considered to be random replication must be taken into account when obtaining the standard errors of the estimated linear and quadratic effects. The standard error for the linear term for the j^{th} fertilizer level is computed from a weighted average of two mean squares, $\text{MS}^*_{z_1} = \frac{1}{t}(\text{MSBlock}_{z_1} + (t\text{-}1)\ \text{MSB}\times\text{F}_{z_1})$, where t is the number of fertilizer treatments. The standard error is $s(\bar{z}_{.j1}) = \sqrt{\frac{\text{MS}^*_{z_1}}{n_{.j}}}$, where $n_{.j}$ is the number of plots receiving the j^{th} fertilizer treatment. The mean values of z_{ij0}, z_{ij1}, and z_{ij2} for the mean of all blocks and all fertilizer treatments, $\bar{z}_{..0}$, $\bar{z}_{..1}$, and $\bar{z}_{..2}$, respectively, provide estimates of the mean, linear, and quadratic effects of time averaged over all fertilizer levels. The standard error of $\bar{z}_{..k}$ is $s(\bar{z}_{..k}) = \sqrt{\frac{\text{MS Block}_{z_k}}{n_{..}}}$, where $n_{..}$ is the total number of plots in the experiment. The estimated mean, linear, and quadratic effects for each nitrogen and phosphorus level, along with their standard errors are given in Table 4.

The intercept and slope in the quadratic regression equation relating tree volume and time both increased as N increased. The effect of phosphorus on intercept and slope was not linear. Mean tree volume and the rate of increase of tree volume were greatest at 28 kg P/ha. The fitted curve for the treatment corresponding to the highest level of N and the middle level of P (the 11^{th} fertilizer level), which gave the maximum response was:

$$\hat{y}_{.11k} = \underset{(223.1)}{5099} \cdot z_0 + \underset{(36.67)}{1091} \cdot z_1 + \underset{(10.57)}{13.19} \cdot z_2.$$

Standard errors of the regression coefficients are in parentheses. Since there were no significant $N \times P$ interactions, each of the coefficients was obtained by adding together the grand mean, the N=336 effect, and the P=28 effect from Table 3; for example, the first term is 5099=4703.56 + (5014.86444 – 4703.56) + (4787.82144 – 4703.56). The standard error is the square root of $\mathrm{Var}(\hat{z}_{.11k}) = \frac{1}{4}(\sigma^2_{\beta zk} + \frac{1}{2}\sigma^2_{\epsilon zk})$, which is estimated by $\frac{1}{48}(\mathrm{MSBlock}_{zk} + 5 \cdot \mathrm{MSB} \times \mathrm{F}_{zk})$. In general, for a randomized block design with r blocks and two between-plot treatment factors, A with a levels and B with b levels, and no significant $A \times B$ interaction, the variance of a fitted value is $\frac{1}{r}\,\sigma^2_{\beta zk} + \frac{a+b-1}{rab}\,\sigma^2_{\epsilon zk}$ (for details see Appendix IV). In terms of years from fertilizer application, the model for predicting tree volume is

$$\hat{y}_{.11k} = 5099 \cdot (\tfrac{1}{\sqrt{3}}) + 1091 \cdot (\tfrac{1}{\sqrt{2}}) \cdot (k-2) + 13.19 \cdot (\tfrac{3}{\sqrt{6}}) \cdot ((k-2)^2 - \tfrac{2}{3}), \text{ where } k = \tfrac{\mathrm{yr}}{2}.$$

Formulas for orthogonal polynomials for equally spaced times are given in Draper and Smith (1981). If the levels of the repeated factor are not equally spaced, the conversion to orthogonal polynomials and back can be done using a Gram-Schmidt orthogonalization program, such as the GSORTH function in SAS® PROC IML. Alternatively, the whole analysis could be done using ordinary polynomials rather than orthogonal polynomials. In this case, the variables $z_{ij0},..,z_{ijp-1}$ would be the regression coefficients for the i,j^{th} plot obtained by fitting a separate polynomial in time for each plot. For a high order polynomial, rounding errors may be a problem and it may not be computationally feasible to fit an ordinary polynomial. In such a case orthogonal polynomials must be used.

The procedure of computing separate regression coefficients for each plot which are then subjected to further analysis produces efficient estimates of the coefficients if there are a total of p independent z variables; i.e., if the model for the repeated factor is saturated. For the case of a polynomial of order less than p-1, no single estimator is most efficient and several different estimators have been proposed (Timm 1980). The procedure just described has the advantage that its exact distribution is known, so that the significance levels of hypothesis tests are exact, the parameter estimates are unbiased, and the standard errors given above are exact. It has the additional advantage that the procedure is relatively simple to implement.

Split plot design with repeated measures

Experimental situation

Here we consider the standard split plot design with main plots in complete blocks. Levels of the main plot factor A are assigned randomly to main plots within blocks, and levels of the subplot factor B are assigned randomly to subplots within each main plot, (i.e., there is no "stripping" of levels of either A or B). In every subplot, measurements are taken at each of p times (or at each of p locations). In contrast to the levels of factors A and B, levels of the repeated measures factor, time or location, are not allocated randomly.

A study on the effects of ozone concentration and pH of rain on the growth of seedlings of three half-sib families of loblolly pine provides an example of a split plot with repeated measures across time (Kress et al. 1988). For illustration, we have utilized a subset of the treatments corresponding to a complete factorial set. Main plots were large open top chambers arranged in 2 complete blocks, 10 chambers per block. The main plot treatments had a factorial structure corresponding to the 10 combinations of 5 ozone (O_3) levels (CF=charcoal filtered air, NF=non-filtered, 1.5x=1.5 times NF conc., 2.25x, and 3.0x) and "rain" at 2 pH levels (pH 3.5 and 5.2).

Family (Fam), with 3 levels, was the subplot factor B, and within each chamber each family was assigned randomly to one of 3 sectors. There were 5 plants per family in each sector or subplot (300 plants in all). Measurements of various aspects of growth were taken on each plant on 12 occasions during the second year after plots were established. For illustration, we will analyze the increase in total height (relative to height before O_3 exposures began) determined on 6 occasions about a month apart and averaged over the 5 plants in a subplot. Thus for each subplot there are 6 growth measurements, and the repeated measures factor is date, with levels corresponding to 29, 57, 85, 113, 156 and 183 days after 3.25.87. For space reasons, the complete data set is not presented.

Assumptions and model

We assume that the main and subplot factors represent fixed effects, but block effects are random. Time main effects are fixed but time × block interaction effects are of course random. Again the systematic nature of the repeated measures factor, time, makes it necessary to allow for possibly different correlations between measurements over time within the same subplot. Similarly, the "equal correlations" assumption may be incorrect for observations across time within the same main plot, or within the same block. Measurements in different blocks are uncorrelated. The resulting model is presented in detail for completeness, but the detail could be skipped at first reading.

Without using vector notation, the linear model is

$$[3] \quad Y_{ijkl} = \mu + \beta_i + A_j + \epsilon_{ij} + B_k + (AB)_{jk} + \delta_{ijk} + T_\ell + (\beta T)_{i\ell} + (TA)_{j\ell} + \theta_{ij\ell} + (TB)_{k\ell} + (TAB)_{jk\ell} + \gamma_{ijk\ell}$$

The A_j, B_k and $(AB)_{jk}$ are fixed main and interaction effects for the factors A and B; and T_ℓ, $(TA)_{j\ell}$, $(TB)_{k\ell}$ and $(TAB)_{jk\ell}$ are a fixed time (or location) main effect and corresponding fixed interaction effects for the repeated measures factor. The β_i, ϵ_{ij}, and δ_{ijk} are random (main) effects for blocks, main plots, and sub-plots respectively, all mutually independent with $Var(\beta_i) = \sigma^2_\beta$,

$\mathrm{Var}(\epsilon_{ij}) = \sigma_\epsilon^2$ and $\mathrm{Var}(\delta_{ijk}) = \sigma_\delta^2$. The $(\beta \mathrm{T})_{i\ell}$, $\theta_{ij\ell}$, and $\gamma_{ijk\ell}$, represent random time-specific contributions for blocks, main plots, and sub plots, respectively, with

$$\mathrm{Cov}((\beta \mathrm{T})_{i\ell}, (\beta \mathrm{T})_{i^*\ell^*}) = \begin{cases} 0 \text{ if } i \neq i^* \\ \sigma_{\beta T \ell \ell^*} \text{ if } i = i^* \end{cases},$$

$$\mathrm{Cov}(\theta_{ij\ell}, \theta_{i^*j^*\ell^*}) = \begin{cases} 0 \text{ if } ij \neq i^*j^* \\ \sigma_{\theta \ell \ell^*} \text{ if } ij = i^*j^* \end{cases},$$

and

$$\mathrm{Cov}(\gamma_{ijk\ell}, \gamma_{i^*j^*k^*\ell^*}) = \begin{cases} 0 \text{ if } ijk \neq i^*j^*k^* \\ \sigma_{\gamma \ell \ell^*} \text{ if } ijk = i^*j^*k^* \end{cases}.$$

Now let $\mathbf{y}_{ijk}$ represent the row vector of p repeated measurements for the ijk^{th} subplot. For example, $\underset{\sim}{\mathbf{y}}_{ijk} = (y_{ijk1}, y_{ijk2}, \cdots, y_{ijk6})$ for a plant in the ozone study. The more compact vector representation of the model in [3] is then

[4] $\quad \underset{\sim}{\mathbf{y}}_{ijk} = \underset{\sim}{\mu} + \underset{\sim}{\beta}_i + \underset{\sim}{\mathbf{A}}_j + \underset{\sim}{\epsilon}_{ij} + \underset{\sim}{\mathbf{B}}_k + (\underset{\sim}{\mathbf{A}}\underset{\sim}{\mathbf{B}})_{jk} + \underset{\sim}{\delta}_{ijk}$

where each of the vectors has p elements, corresponding to each of the p measurement times, and

$\mathrm{Var}(\underset{\sim}{\mathbf{y}}_{ijk}) = \underset{\sim}{\Sigma}_\beta + \underset{\sim}{\Sigma}_\epsilon + \underset{\sim}{\Sigma}_\delta$	for observations in the same sub-plot,
$\mathrm{Cov}(\underset{\sim}{\mathbf{y}}_{ijk}, \underset{\sim}{\mathbf{y}}_{ijk^*}) = \underset{\sim}{\Sigma}_\beta + \underset{\sim}{\Sigma}_\epsilon$	for different sub-plots in the same main plot,
$\mathrm{Cov}(\underset{\sim}{\mathbf{y}}_{ijk}, \underset{\sim}{\mathbf{y}}_{ij^*k^*}) = \underset{\sim}{\Sigma}_\beta$	for different main plots in the same block,
$\mathrm{Cov}(\underset{\sim}{\mathbf{y}}_{ijk}, \underset{\sim}{\mathbf{y}}_{i^*j^*k^*}) = \underset{\sim}{0}$	for plots in different blocks.

The covariance structure for the model in [4] is related to that in [3] as follows: the $\ell\ell^{*th}$ element (relating to measurements at times ℓ and ℓ^*) of $\underset{\sim}{\Sigma}_\beta$ is $\sigma_\epsilon^2 + \sigma_{(\beta T)\ell\ell^*}$, element $\ell\ell^*$ of $\underset{\sim}{\Sigma}_\epsilon$ is $\sigma_\epsilon^2 + \sigma_{\theta\ell\ell^*}$, and element $\ell\ell^*$ of $\underset{\sim}{\Sigma}_\delta$ is $\sigma_\delta^2 + \sigma_{\gamma\ell\ell^*}$. Although variances and covariances are allowed to be time specific (i.e., dependent on the levels of the repeated measures factor) models [3] and [4] imply a "homogeneity of variance" assumption in that $\underset{\sim}{\Sigma}_\beta$, $\underset{\sim}{\Sigma}_\epsilon$ and $\underset{\sim}{\Sigma}_\delta$ are each constant across "treatments" (i.e., the same for all combinations of levels of A and B).

Analysis: hypothesis and test procedures

Except for the need to carry out additional tests for effects involving the subplot factor, the analysis for the split plot design with repeated measures is analogous to that described for the randomized complete block design. We therefore omit most of the detail and illustrate the main features of the analysis using the O_3 - acid rain study as an example.

Again it is useful to think of the analysis as consisting of two parts, a between-plot analysis and a within-plot analysis. The between-plot analysis is a split plot analysis applied to the subplot means computed by averaging over all 6 levels of the repeated measures factor, date. In the between-plot analysis, MANOVA and the "analysis of contrasts" yield identical tests, and all tests relate to effects of the main plot and subplot factors averaged over dates. The within-plot analysis provides tests for all effects involving the repeated measures factor, date. Omnibus multivariate tests are carried out, as well as tests aimed at more specific aspects of the repeated measures factor. These latter tests are usually based on using orthogonal contrasts to partition the effect of the repeated measures factor, resulting in a series of $p - 1$ separate split plot analyses, each on a different contrast. Interpretation of the omnibus multivariate tests and the contrast - based tests is given below.

Plotting growth against date for the various O_3-acid rain treatments (Fig. 1) suggested that a simple polynomial equation would describe the response curve, and so orthogonal polynomial contrasts were used to partition the date effect. Results for the between-plot analysis (on the means over dates) and sums of squares for the $p-1=5$ polynomial contrasts are presented in Table 5. Note that for simplicity the factorial structure of the main plot treatments has been suppressed, the main plot factor being denoted "Trt", the subplot factor being "Fam".

All tests in the between-plot analysis relate to effects that are averaged over dates. Thus the test for Trt tests the null hypothesis that mean growth (averaging over the 6 dates and over the 3 families) is the same for each of the 10 $O_3 \times pH$ combinations. In the notation of model [3], the null hypothesis is H_0: $A_1 = A_2 = \ldots = A_{10}$, where A_i is the main effect for the i^{th} $O_3 \times pH$

combination. Similarly, the test for Fam is a test for no family main effect averaging both over the 10 Trts and over the 6 dates. In the notation of [3], the null hypothesis is H_0: $B_1 = B_2 = B_3$, where B_i is the main effect for the i^{th} family. Only the Fam main effect appears important, but before concluding that Trt effects are nonsignificant it is necessary to examine results for the interactions with date in the within-plot part of the analysis.

In the within-plot analysis, the MANOVA test for a date main effect would test that mean height (averaging over treatments and families) is the same at each date. Alternatively, this test can be interpreted as testing simultaneously that the linear, quadratic, up through quintic, components of the regression of height on date are each zero. The analysis of contrasts differs from MANOVA in that a separate test is performed for each of the polynomial contrasts. In the O_3-acid rain example, the MANOVA test for date cannot be carried out as there are only two blocks, yet the main effect for each of the five polynomial contrasts can be tested separately via the analysis of contrasts. In situations like this where the number of blocks is less than the number of measurement times, we suggest limiting attention to two or three contrasts (selected a priori) as a compromise between not performing a test for a date main effect and on the other hand possibly incurring an unsatisfactory type I error rate because of multiple testing (see also Eskridge and Stevens 1987).

For studies like the present example, where growth is measured over weeks or months, it will usually be obvious that there is a time main effect, and an omnibus test for such an effect seems superfluous. Certainly for our example, demonstrating a date main effect statistically seems unimportant as there is no doubt that the plants have grown during the six month period of observation. Of greater interest is describing the growth curve and how this curve is affected by the main and subplot factors. As can be seen from the sums of squares in Table 5, the linear and quadratic components of the regression of height on date explain most of the variation due to date. The effects of the treatments on growth can therefore be examined by focusing on the interactions between the linear and quadratic polynomial contrasts and the main and subplot factors. This is consistent with our recommendation to limit attention to only a few contrasts when the number of blocks is less than the number of levels of the repeated measures factor.

Results for the tests carried out in the within-plot analysis are presented in Table 6 for the linear and quadratic contrasts. The *p*-values for the omnibus MANOVA tests (excepting, of course, the test for a date main effect) are also given in Table 6. Note that the factorial structure of the main plot treatments is recognized in Table 6 and the effects for Trt partitioned into pH, O_3, and pH $\times$ O_3 components. Also indicated in Table 6 are the error terms for each of the F tests in the analysis of each contrast, and the analogous error matrix used in the corresponding MANOVA test. As noted above, F tests for each contrast are conducted as for a split plot design with main plots in blocks, blocks assumed random, and with a test for the grand mean being zero included.

The MANOVA tests suggest that there are no strong interactions between date and any of the main or subplot factors, i.e., that the growth curves are similar across treatments. If a 10% significance level is used then there are indications of possible differences in growth among the combinations of pH, O_3, and Fam (p=.093 for Date $\times$ pH, p=.076 for Date $\times$ Fam, and p=.047 for Date $\times$ O_3 $\times$ Fam).

Compared to the multivariate tests, those carried out in the analysis of contrasts can have greater power for detecting treatment $\times$ time interactions if the effect of date or time can be summarized in one or two contrasts. This is analogous to the increase in power for specific hypotheses obtained by partitioning a treatment sum of squares into components representing, for example, factorial main and interaction effects. Results for the analysis of the linear and quadratic contrasts on date are therefore of interest, even if the MANOVA tests are largely nonsignificant.

For the date linear contrast, the test for the mean (i.e. date effect) tests the null hypothesis that, averaging over all levels of Trt (i.e., pH $\times$ O_3 combinations) and Fam, the regression of growth (equivalently height) on date has no linear component, Roughly, this corresponds to testing that the regression of growth on date (after averaging over Trt and Fam) has slope 0. For the linear contrast, the Date*pH effect provides a test of whether, averaging over O_3 levels and families, the regression of growth on date has the same "slope" or linear component for the two pH levels. Similarly, the test of Date*Fam tests the null hypothesis that, averaging over all 10 levels of Trt, the regression of growth on date has the same slope for each of the 3 families. The Date*pH*Fam interaction tests whether (averaging over O_3 levels) the difference between slopes for the two pH levels is the same for each of the 3 families.

Interpretation of the tests for the quadratic contrast is the same except that "quadratic component" is substituted for "slope" or "linear component". There appear to be differences in the linear components of the regressions for families and these differences depend on pH ($p = .050$ and $p = .046$ for the Fam and pH × Fam effects, respectively). A slightly stronger effect is that the curvature of the regressions for the 2 pH levels seems to be different ($p = .001$ for pH in the analysis of the quadratic contrast). Plotting mean growth for the 6 pH × Fam combinations against date shows that growth appears to be slightly greater for rain at pH 5.2 than for pH 3.5 in families 2 and 3, but the opposite is true in the last two months for family 1. As noted in Kress, et al. (1988), data from subsequent years should help to determine whether this represents a real effect of the rain acidity or not.

Neither the linear nor the quadratic contrast on date indicates a significant effect of O_3 concentration on growth, although Kress et al. (1988) note that growth was suppressed at the two highest O_3 levels. The O_3 interaction effects could be examined more closely by partitioning them into appropriate contrasts in Table 6, but we have chosen instead to keep the analysis as simple as possible.

Finally, we comment on an approach that is often incorrectly used for the analysis of data such as the O_3-acid rain growth data. This approach is to carry out a split-split-plot analysis, treating the repeated measures factor date as a sub-sub-plot factor, and results for effects involving date from the sub-sub-plot part of this analysis are given in Table 6 for comparison with the MANOVA and analysis of contrasts. Validity of these sub-sub-plot tests requires homogeneity of variances and covariances across dates, an assumption that often fails for repeated measurements on height, not only because of the systematic nature of the factor time or date, but also because variance tends to increase as mean height increases. Evidence that the variance-covariance structure is not independent of date is seen in the error sums of squares (i.e. $\text{Date} \times \text{Bl} \times \text{Fam(Trt)}$) for the polynomial contrasts in Table 5. These sums of squares should be of about the same magnitude under homogeneity of variances and covariances, but are seen to differ by up to 2 orders of magnitude. For repeated measures data, dependence on time of the variance-covariance structure will tend to result in liberal tests, that is, tests which produce too many significant outcomes (e.g. Meredith and Stehman, 1991). From Table 6, we see that the agreement between p-values produced by the split-split-plot tests and those produced by the MANOVA tests is poor, even though the hypotheses examined are equivalent. The tests based on the linear and quadratic contrasts examine more specific hypotheses and so p-values need not agree even qualitatively with those from the split-split-plot analysis or MANOVA. Generally, however, there is a tendency for the split-split-plot tests to suggest the presence of stronger treatment and family effects on growth than either the MANOVA or analysis of contrasts. As plots do not indicate strong effects on growth (see Figure 1) this illustrates that the split-split-plot p-values can be misleading.

Discussion and conclusions

For experiments with treatments in a blocked design and repeated measures over time or location, all of the main effects and interactions can be tested in a multivariate analysis of variance framework. In addition to these omnibus tests for the repeated factor and its interactions, it is usually of interest to study specific contrasts of the repeated factor. Commonly used contrasts include the differences between adjacent times, the difference between each time and the initial time, polynomial contrasts, and comparison of each time to the average of the remaining times to determine when a plateau has been reached. The multivariate analysis provides nonspecific tests concerning the presence of treatment or time effects, while the analysis of contrasts allows more detailed exploration of the nature of the response over time and how this is affected by the treatments. The analysis of contrasts is more flexible than the multivariate analysis, so sometimes only part of the full analysis can be completed. For example, in randomized block designs it often happens that there are more levels of the repeated factor than there are blocks. In this case it is not possible to compute the multivariate test statistic for the repeated factor main effect. It is, however, still possible to test the individual contrasts of the repeated factor. Even when there are only a few levels of the repeated factor, if there are only three or four blocks the multivariate test may tend to lack power (Davidson 1972). When the number of blocks is small, the analysis of contrasts may prove more informative than the multivariate tests. In standard software such as the SAS PROC GLM "REPEATED" statement the multivariate tests require a complete set of observations for every plot. If a few data points are missing from one or more plots, it is still possible to do the analysis of contrasts, using PROC GLM without the "REPEATED" statement.

There are some situations for which the methods of this paper are not adequate. If more than a few data points are missing or if the levels of the repeated factor are different for different plots then the multivariate analysis is impossible and the analysis of contrasts may be less efficient than methods such as estimated generalized least squares. Covariates that vary from plot to plot can be incorporated into the analyses presented in this paper, but covariates that vary over time, such as rainfall between measurement times at different sites, cannot. In these circumstances more sophisticated methods such as estimated generalized least squares or restricted maximum likelihood should be used rather than the repeated measures analysis presented here (Schaalje et al. 1991, McLean et al. 1991). Growth curves are usually s-shaped, with growth approaching an asymptote, and are not fit very well by polynomials. If the aim is to describe and predict growth by fitting a mathematical model, then a nonlinear model with an appropriate covariance structure accounting for correlations over time should be used (Lindstrom and Bates 1990, Vonesh and Carter 1991).

The models we have discussed are appropriate for the most commonly used agricultural experimental designs which, in our experience, are factorials in randomized complete blocks and split plots or split-split plots. Taking several observations on each experimental unit is an efficient way to obtain more information from a fixed amount of experimental material. When planning a repeated measures experiment, the number of blocks should exceed the number of contrasts needed to describe the effect of the repeated factor. In order to compute all of the multivariate test statistics at least as many blocks are required as there are levels of the repeated factor. It is reasonable to require more experimental units in a repeated measures design because each experimental unit provides information on a possibly complex time (or location) effect. Muller and Peterson (1984) give methods for computing the power of multivariate tests which can be used in planning sample sizes.

Some texts recommend randomizing the order of treatments on each experimental unit (e.g. Johnson and Wichern 1988). In forestry the most commonly used repeated measures factors, such as time, depth, and distance from the row, cannot be randomly allocated. If the experimental unit is the tree and different treatments are applied to different leaves or branches of each tree, the treatments should be randomly allocated. Cases where different treatments are assigned to each experimental unit at different times are more complicated. Examples include crop rotation studies and studies involving human subjects. For instance, in a study of the efficacy of three different remedies for chronic pain, each patient might receive all three drugs given in random order, spaced out at suitable time intervals. In this type of design there are possible complications of order effects and carry-over effects from one time period to the next and there is the additional problem that it is difficult to know whether correlation patterns should be caused by the passage of time or by the individual's type of response to different types of treatments. That is, should the repeated factor be "drug" or should it be "time" for purposes of estimating the covariance matrices?

We have demonstrated repeated measures analysis for a randomized complete block design and for a split plot experiment in which the repeated factor is time and all plots are measured at the same times. The methods described in this paper can be extended to other balanced designs with random effects that incorporate repeated measures. Factors that often are considered to be random include sites and families. The analysis for an experiment with one random effect and one fixed effect differs from the case when both factors are fixed in two respects:

1. The error term for testing the main effect of the fixed factor involves the sum of squares for the interaction of the fixed effect with the random effect. This is true for all univariate and multivariate tests. For example, if A is fixed and B is random, then the univariate and multivariate test statistics are $F = MSA/(MSA \times B)$, and $\Lambda = |\underset{\sim}{SSCP}(A \times B)| / |\underset{\sim}{SSCP}(A) + \underset{\sim}{SSCP}(A \times B)|$, respectively. The proper tests may not be provided automatically by standard statistical software and it may be necessary to write a few extra commands to compute them.

2. Standard errors of treatment means and regression coefficients incorporate the variation among levels of the random factor as well as the within-plot variation. They are not simply computed from MSE. The proper standard errors can be determined from inspection of the univariate expected mean squares. Again, statistical software may not provide the correct standard errors, and some hand computations may be required. Usually the standard errors for treatment means averaged over levels of a random factor will be larger than if all factors had been fixed.

Acknowledgments

We thank H. Lee Allen and Lance Kress for providing the air pollution data set, and the North Carolina State Forest Nutrition Cooperative and Weyerhaeuser Company for use of the mid-rotation fertilizer data.

References

DAVIDSON, M.L. 1972. Univariate versus multivariate tests in repeated-measures experiments. Psychol. Bull. 77: 446-452.

ESKRIDGE, K.M., and STEVENS, E.J. 1987. Growth curve analysis of temperature-dependent phenology models. Agron. J. 79: 291-297.

JOHNSON, D.E., CHAUDHURI, U.N., and KANEMASU, E.T. 1983. Statistical analysis of line-source sprinkler experiments and other nonrandomized experiments using multivariate methods. Soil Sci. Soc. Am. J. 47: 309-312.

JOHNSON, R.A., and WICHERN, D.W. 1988. Applied Multivariate Statistical Analysis, Second Ed. Prentice Hall, Englewood Cliffs, NJ. 607 pp.

KEMPTHORNE, O. 1952. The Design and Analysis of Experiments. Reprinted 1979. Robert E. Krieger Publishing Co. Huntington, NY. 631 pp.

KRESS, L.W., ALLEN, H.L., MUDANO, J.E., and HECK, W.W. 1988. Response of loblolly pine to acidic precipitation and ozone. Proceedings of APCA 81st Ann. Mtg. Dallas, TX.

LINDSTROM , M. J., and BATES, D. M. 1990. Nonlinear mixed effects models for repeated measures data. Biometrics. 46: 673-687.

MCLEAN, R.A., SANDERS, W.L., and STROUP, W.W. 1991. A unified approach to mixed linear models. Am. Stat. 45: 54-64.

MEREDITH, M.P., and STEHMAN, S.V. 1991. Repeated measures experiments in forestry: focus on analysis of response curves. Can. J. For. Res. 21: 957-965.

MORRISON, D.F. 1990. Multivariate Statistical Methods, Third Ed. McGraw-Hill. New York. 495 pp.

MOSER, E. B. , SAXTON, A.M., and PEZESHKI, S.R. 1990. Repeated measures analysis of variance: application to tree research. Can. J. For. Res. 20: 524-535.

MULLER, K.E., and PETERSON, B.L. 1984. Practical methods for computing power in testing the multivariate general linear hypothesis. Computational Stat. & Data Anal. 2: 143-158.

ROWELL, J.G., and WALTERS, D.E. 1976. Analysing data with repeated observations on each experimental unit. J. Agric. Sci. 87: 423-432.

SCHAALJE, B., ZHANG, J., PANTULA, S.G., and POLLOCK, K.H. 1991. Analysis of repeated-measurements data from randomized block experiments. Biometrics 47: 813-824.

TIMM, N.H. 1980. Multivariate analysis of variance of repeated measurements. Handbook of Statistics, Vol. I, P. R. Krishnaiah, ed. North-Holland Publishing Co. pp 41-87.

VALENTINE, D.W., and ALLEN, H.L. 1990. Foliar responses to fertilization identify nutrient limitation in loblolly pine. Can. J. For. Res. 20: 144-151.

VONESH, E.F., and CARTER, R.L. 1991. Mixed effects nonlinear regression of unbalanced repeated measures. Biometrics. In press.

VOSE, J.M., and ALLEN, H.L. 1988. Leaf area, stemwood growth, and nutrition relationships in loblolly pine. For. Sci. 34: 547-563.

Table 1. Tree volume (ft^3/acre) from loblolly pine mid-rotation fertilization experiment.

		Block 1			Block 2			Block 3			Block 4		
N	P	Year			Year			Year			Year		
(kg/ha)	(kg/ha)	2	4	6	2	4	6	2	4	6	2	4	6
	0	1658	2119	2687	1755	2272	2762	2068	2618	3208	2237	2921	3623
0	28	1726	2249	2843	1847	2376	2909	1988	2511	3133	2115	2784	3412
	56	1794	2375	3007	1839	2332	2914	2006	2608	3250	2247	2742	3398
	0	1788	2329	2742	1864	2387	2988	2159	2840	3418	2387	3051	3765
112	28	1989	2575	3242	1976	2615	3250	2160	2837	3532	2106	2698	3318
	56	1805	2424	3024	1850	2447	3053	1999	2618	3192	2181	2777	3498
	0	1875	2502	3054	1913	2556	3169	2149	2893	3685	2409	3155	3847
224	28	1812	2412	3089	1981	2659	3432	2293	3063	3855	2490	3273	4091
	56	1819	2497	3208	2089	2769	3703	2263	3076	3873	2283	2826	3535
	0	1934	2675	3313	1984	2637	3339	2381	3093	3843	2336	3068	3764
336	28	2116	2926	3681	1964	2713	3472	2237	2968	3832	2523	3328	4276
	56	1880	2543	3165	1985	2690	3454	2168	2951	3742	2317	3081	3846

Table 2. Analysis of time contrasts for tree volume measurements from the fertilizer experiment.

Between-plot analysis

Source	df	SS	MS	F	P-value
Mean	1	1061928863	1061928863	490	.0002
Block	3	6503411	2167804	49.0	
(Fertilizer	11)				
N	3	3334687	1111562	25.1	.0001
P	2	170444	85222	1.93	.16
N × P	6	224305	37384	0.84	.55
Block × Fertilizer	33	1460724	44264		

Within-plot analysis

Source	df	SS	MS	F	P-value
Time Linear	1	42264950	42264950	1146	.0001
Linear × Block	3	110624	36875		
Linear × N	3	461468	153823	27.8	.0001
Linear × P	2	47871	23935	4.32	.022
Linear × N × P	6	24029	4005	0.72	.63
Linear × Fert × Block	33	182685	5536		
Time Quadratic	1	7503	7503	6.81	.080
Quadratic × Block	3	3305	1102		
Quadratic × N	3	2720	907	1.06	.38
Quadratic × P	2	6088	3044	3.57	.039
Quadratic × N × P	6	4500	750	0.88	.52
Quadratic × Fert × Block	33	28115	852		

Table 3. Matrices of sums of squares and cross products for a randomized complete block design

Source	$df^{\dagger}$	matrix
Mean	1	$SS\underset{\sim}{C}P(\text{Mean})_{\mathbf{y}}$
Block	b-1	$SS\underset{\sim}{C}P(\text{Block})_{\mathbf{y}}$
A	a-1	$SS\underset{\sim}{C}P(\text{A})_{\mathbf{y}}$
Block × A	(b-1)(a-1)	$SS\underset{\sim}{C}P(\text{Block}\times\text{A})_{\mathbf{y}}$

† Degrees of freedom associated with each element of the matrix. Number of blocks=b, levels of A=a.

Table 4. Mean, linear, and quadratic time effects for levels of N and P in the fertilizer example.

N (kg/ha)	z_0	z_1	z_2
0	4346	817.1	20.82
112	4565	869.6	3.062
224	4888	1011	18.88
3361	5015	1055	7.246
std. error*	218.9	33.38	8.729

$$* \text{ S.E.} = \sqrt{\frac{\text{MSBlock}_{z_k} + 3 \cdot \text{MSB} \times \text{F}_{z_k}}{4 \cdot 12}}.$$

P (kg/ha)	z_0	z_1	z_2
0	4663	897.6	-3.266
28	4788	974.5	18.45
56	4660	943.0	22.33
std. error†	216.8	31.60	7.645

$$\dagger \text{ S.E.} = \sqrt{\frac{\text{MSBlock}_{z_k} + 2 \cdot \text{MSB} \times \text{F}_{z_k}}{3 \cdot 16}}.$$

Table 5. Between plot analysis and within plot sums of squares for the pine seedling growth (cm) data measured on 6 dates in the ozone-acid rain study.

Between-plot analysis

Source	df	SS	F	$P-$value
Block	1	38.94	0.16	0.700
Trt	9	2805.13	1.26	0.367
Block*Trt[a]	9	2219.51	1.91	0.109
Fam	2	1186.78	4.60	0.023
Trt*Fam	18	2092.32	0.90	0.586
Block*Fam(Trt)[b]	20	2582.34		

Within-plot analysis – sums of squares for the 5 orthogonal polynomial contrasts on date

Source	df	Date linear	Date quadratic	Date cubic	Date quartic	Date quintic
Date	1	138293.8	2370.4	335.45	5.64	2.15
Date*Block	1	47.2	92.4	0.74	9.98	2.25
Date*Trt	9	1242.3	130.9	11.28	37.34	4.04
Date*Block*Trt	9	767.6	36.9	34.19	7.36	7.82
Date*Fam	2	239.1	18.3	5.23	0.33	0.89
Date*Trt*Fam	18	699.6	55.6	27.71	24.84	16.17
Date*Block*Fam(Trt)	20	682.7	97.7	20.62	20.74	7.59

[a]Error for testing Block and Trt in the between-plot analysis

[b]Error for testing Fam and Fam*Trt in the between-plot analysis

Table 6. Within-plot analysis of growth data from the ozone-acid rain study. P – values for tests of Date effects based on (i) MANOVA, (ii) analysis of orthogonal polynomial contrasts on Date, and (iii) an invalid split-split-plot analysis, treating date as a sub-sub-plot factor.

	Analysis Method			
	(i) MANOVA	(ii) Analysis of Contrasts[a]		(iii) Split-split plot
Source		linear	quadratic	
Date[b]		0.012	0.124	< 0.0001
Error 1[c]	0.100	0.476	0.001	0.177
Date*pH[c]	0.093	0.851	0.001	0.351
Date*O_3[c]	0.835	0.278	0.188	0.136
Date*pH*O_3	0.537	0.160	0.579	0.028
Error 2[d]	0.024	0.042	0.589	0.0003
Date*Fam[d]	0.129	0.050	0.179	0.001
Date*Fam*pH[d]	0.076	0.046	0.884	0.002
Date*Fam*O_3[d]	0.047	0.324	0.385	0.165
Date*Fam*pH*O_3	0.547	0.898	0.971	0.998
Error 3 MS[e]		34.13	4.88	8.293
(df)		(20)	(20)	(100)

[a]Polynomial contrasts of degree greater than 2 accounted for only a small fraction of the variation

[b]Indicates effect tested against Error 1 (Date*Block Error)

[c]Indicates effect tested against Error 2 (Date*Main plot Error)

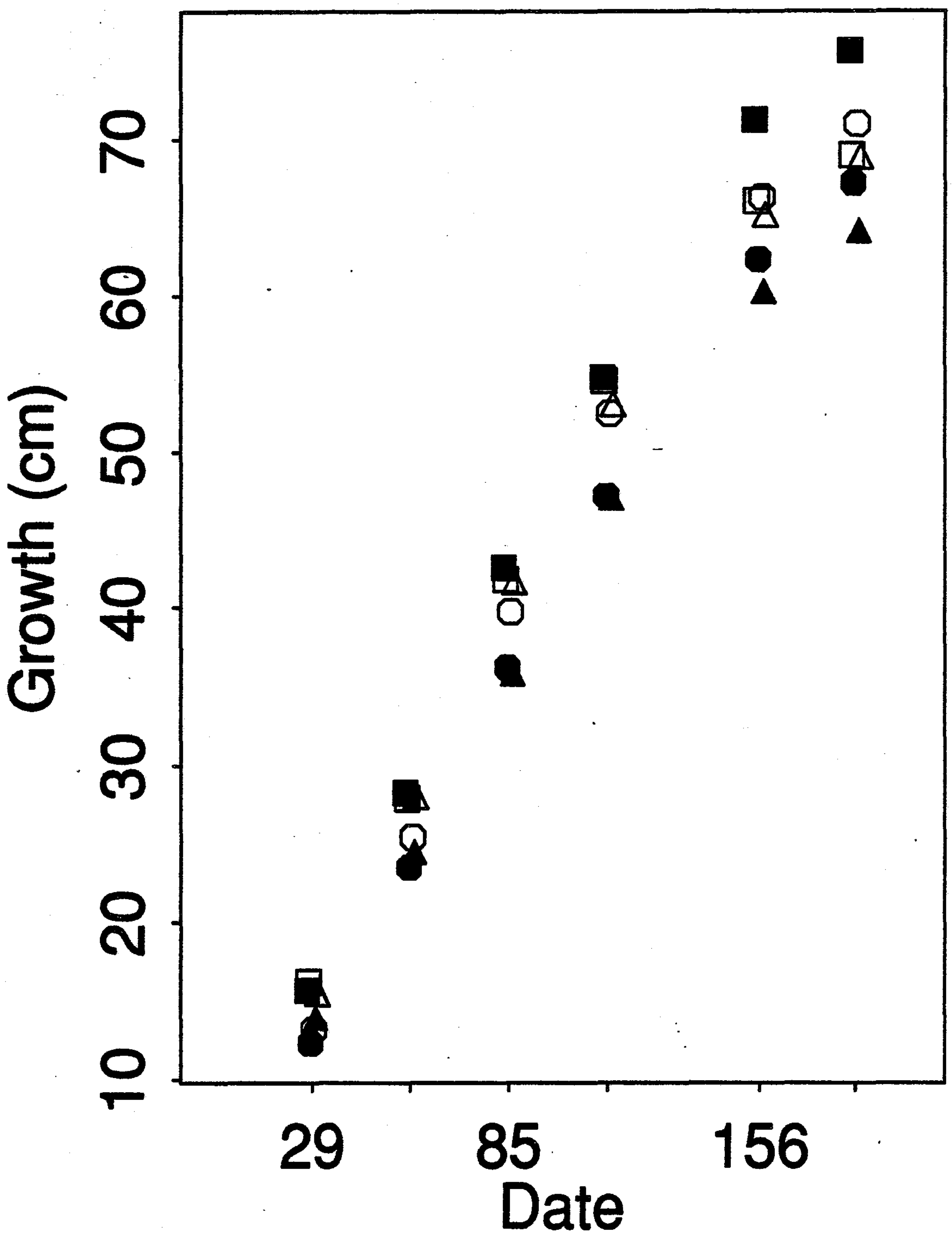
[d]Indicates effect tested against Error 3 (Date*Sub plot Error)

[e]For univariate analyses only

Figure Legend

Figure 1. Mean growth (increase in height from 3.25.87) measured at 6 dates after 3.25.87 for 3 families of loblolly pine exposed to "rain" at pH 5.2 or at pH 3.5. Shaded symbols (□, ○, and △) represent means for families 1,2, and 3 respectively, for rain at pH 3.5, and open symbols (□, ○, and △) represent means for families 1,2, and 3, respectively, for rain at pH 5.2.

Figure 1.



Appendix I. SAS commands for randomized blocks with repeated measures

The SAS statements for producing the multivariate tests and the analysis of polynomial contrasts are given below. The "REPEATED" statement provides tests for N, P, and time main effects and their interactions, but uses the residual error matrix, **SSCP(Block × F)**, for all tests. This is appropriate for tests involving the between-plot factors N and P, but not for the time main effect. The "MANOVA" statement is included to compute the tests for the time effects using **SSCP(Block)** as the error matrix. The "SUMMARY" option on the "REPEATED" and "MANOVA" statements provides the analysis of variance tables for the linear and quadratic contrasts.

```
PROC GLM DATA=LOBLOLLY;
  CLASS BLOCK N P;
  MODEL Y1 Y2 Y3 = BLOCK N P N*P / NOUNI;
  REPEATED TIME 3 POLYNOMIAL / PRINTH PRINTE SUMMARY;
  MANOVA H=INTERCEPT E=BLOCK M=(1 1 1) MNAMES=M0/ ORTH SUMMARY;
  MANOVA H=INTERCEPT E=BLOCK M=(-1 0 1, 1 -2 1) MNAMES=LINEAR QUAD /
    PRINTH PRINTE SUMMARY ORTH;
```

The next set of SAS statements produces the estimated values of the contrasts. The "LSMEANS" statement produces means of the regression coefficients for different levels of N and P. The fitted value for a particular fertilizer level (assuming there is no N × P interaction) can be obtained using the "ESTIMATE" statement as in the example (N=336, P=28), but the standard error must be computed by hand. The standard error printed out by the "ESTIMATE" statement is incorrect when there are random block effects.

```
DATA Z;
  SET LOBLOLLY;
  Z0=(Y1+Y2+Y3)/SQRT(3);
  Z1=(Y3-Y1)/SQRT(2);
  Z2=(Y1 – 2*Y2 + Y3)/SQRT(6);
PROC GLM DATA=Z;
  CLASS BLOCK N P;
  MODEL Z0 Z1 Z2 = BLOCK N P N*P ;
  LSMEANS N P;
  ESTIMATE 'N336 P28'   INTERCEPT 12   BLOCK 3 3 3 3   N 0 0 0 12   P 0 12 0
                    N*P -1 2 -1 -1 2 -1 -1 2 -1 3 6 3   /  DIVISOR=12;
```

Appendix II. $\underset{\sim}{\mathbf{L}}$ and $\underset{\sim}{\mathbf{M}}$ matrices for the fertilizer experiment.

Factor	Contrast Matrix
Mean (across plots)	$\underset{\sim}{\mathbf{L}}_0 = \frac{1}{8\sqrt{3}}\begin{bmatrix} 12 & 3 & 3 & 3 & 3 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$
Nitrogen	$\underset{\sim}{\mathbf{L}}_n = \frac{1}{\sqrt{6}}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & -1 & -1 & -1 \end{bmatrix}$
Phosphorus	$\underset{\sim}{\mathbf{L}}_p = \frac{1}{\sqrt{8}}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 1 & 0 & -1 & 1 & 0 & -1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 \end{bmatrix}$
N × P	$\underset{\sim}{\mathbf{L}}_{np} = \frac{1}{2}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & -1 & 1 \end{bmatrix}$
Mean (across times)	$\underset{\sim}{\mathbf{M}}_0 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$
Time	$\underset{\sim}{\mathbf{M}}_t = \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{6} \\ 0 & -2/\sqrt{6} \\ 1/\sqrt{2} & 1/\sqrt{6} \end{bmatrix}$

Any set of p-1 linearly independent contrasts of the repeated factor could be used in the $\underset{\sim}{\mathbf{M}}_t$ matrix since the multivariate tests for time main effect and time × fertilizer interactions are invariant to the particular choice of contrasts. All contrasts in the $\underset{\sim}{\mathbf{M}}$ matrices have been scaled so that each column has length one, i.e. the coefficients squared sum to one in each column.

Appendix III. Test Statistics and Critical Values for Wilks' Lambda

Reject H_0 if the test statistic exceeds the critical value (Morrison 1990).

Case	Test Statistic	Critical Value
$\min(d_h, c)=1$	$\frac{d_e - c + 1}{\lvert d_h - c\rvert + 1}\left(\frac{1-\Lambda}{\Lambda}\right)$	$F_{1-\alpha}(\lvert d_h - c\rvert + 1,\ d_e - c + 1)$
$\min(d_h, c)=2$	$\frac{d_e - c + 1}{\lvert d_h - c\rvert + 2}\left(\frac{1-\sqrt{\Lambda}}{\sqrt{\Lambda}}\right)$	$F_{1-\alpha}(2\lvert d_h - c\rvert + 4,\ 2(d_e - c + 1))$
$\min(d_h, c) \geq 3$ if d_e is large	$-(d_e - \frac{1}{2}(c - d_h + 1)) \ln \Lambda$	$\chi^2_{1-\alpha}(c \cdot d_h)$

Symbols in the table:

$c = p - 1$, the number of contrasts of the within-plot factor

$d_h =$ hypothesis degrees of freedom (For a factor, A, with a levels, $d_h = a - 1$)

$d_e =$ error degrees of freedom (for a randomized block design with b random blocks and one fixed treatment factor, A, $d_e = (a-1)(b-1)$ for testing time × A interactions, and $d_e = b - 1$ for testing the time main effect.

Appendix IV. Fitted values and standard errors of coefficients

For a randomized block design with two treatment factors, F and G, we first fit a treatment means model, $z_{ij} = \mu + \beta_i + A_j + \epsilon_{ij}$, where A_j indicates the j^{th} treatment combination (which is a combination of factors F and G). If there is no significant F × G interaction then the fitted coefficients, $\hat{z}_{.j}$, for the j^{th} treatment are estimated using an additive model and switching to double subscripts to denote the levels of F and G (where i=1,...,b blocks, ℓ=1,...,f levels of F, m=1,...,g levels of G) gives

$$\hat{z}_{.j} = \hat{z}_{.\ell m} = \hat{\mu} + \hat{F}_\ell + \hat{G}_m,$$

$$= \bar{z}_{...} + (\bar{z}_{.\ell.} - \bar{z}_{...}) + (\bar{z}_{..m} - \bar{z}_{...})$$

$$= \frac{f+g-1}{bfg} \sum_{i=1}^{b} z_{i\ell m} + \frac{f-1}{bfg} \sum_{i=1}^{b} \sum_{m' \neq m} z_{i\ell m'} + \frac{g-1}{bfg} \sum_{i=1}^{b} \sum_{\ell' \neq \ell} z_{i\ell' m}$$

$$- \frac{1}{bfg} \sum_{i=1}^{b} \sum_{\ell' \neq \ell} \sum_{m' \neq m} z_{i\ell' m'}.$$

Writing each z_{ij} in terms of the model elements yields the variance expression,

$$\mathrm{Var}(\hat{z}_{.j}) = \frac{1}{b}\sigma_\beta^2 + \frac{f+g-1}{bfg}\sigma_\epsilon^2.$$

The variance of any linear combination of coefficients can be found by working with the matrix formulation of the repeated measures model,

$$\underset{\sim}{y} = \begin{bmatrix} \underset{\sim}{y}_{11} \\ \vdots \\ \underset{\sim}{y}_{ba} \end{bmatrix} = \underset{\sim}{X}\,\underset{\sim}{\beta} + \underset{\sim}{\epsilon},$$ where there are b blocks and a treatments.

To describe the effect of the repeated factor using a polynomial model, express $\underset{\sim}{\beta}$ as $\underset{\sim}{\gamma}\,\underset{\sim}{M}'$, where $\underset{\sim}{M}$ is the model matrix for the repeated factor. In the forest nutrition experiment the matrix of regression coefficients, $\underset{\sim}{\gamma}$, and the transposed $\underset{\sim}{M}$ matrix for fitting a second order polynomial to p=3 times would be

$$\underset{\sim}{\gamma} = \begin{bmatrix} \underset{\sim}{\gamma}_{\mu} \\ \underset{\sim}{\gamma}_{block1} \\ \vdots \\ \underset{\sim}{\gamma}_{tmt1} \\ \vdots \\ \underset{\sim}{\gamma}_{tmt12} \end{bmatrix} \text{ and } \underset{\sim}{M}' = \begin{bmatrix} 1 & 1 & 1 \\ t_1 & t_2 & t_3 \\ t_1^2 & t_2^2 & t_3^2 \end{bmatrix}.$$

The elements of $\underset{\sim}{\gamma}$ are the regression coefficients (for the regression on time) for the treatment groups defined in the $\underset{\sim}{X}$ matrix. The row vectors $\underset{\sim}{\gamma}_{effect}$ each contain an intercept, a slope, and a quadratic coefficient; e.g., $\underset{\sim}{\gamma}_{\mu} = [\gamma_{\mu 1}\gamma_{\mu 2}\gamma_{\mu 3}]$. Our interest lies in estimating $\underset{\sim}{\gamma}$ and linear combinations of the elements of $\underset{\sim}{\gamma}$.

Incorporating the model for the time effects into the repeated measures model gives:

$$\underset{\sim}{y}' = \underset{\sim}{M}\underset{\sim}{\gamma}'\underset{\sim}{X}' + \underset{\sim}{\epsilon}'$$

$$\text{vec}(\underset{\sim}{y}') = (\underset{\sim}{X} \otimes \underset{\sim}{M})\ \text{vec}(\underset{\sim}{\gamma}') + \text{vec}(\underset{\sim}{\epsilon}').$$

$\text{Vec}(\underset{\sim}{y}') = [\underset{\sim}{y}_{11} \cdots \underset{\sim}{y}_{ba}]'$ contains the observation vectors for all $n_{..} = ba$ of the plots stacked on top of each other. $\text{Vec}(\underset{\sim}{\gamma}') = [\underset{\sim}{\gamma}_{\mu}\ \underset{\sim}{\gamma}_{block1} \cdots \underset{\sim}{\gamma}_{blockb}\ \underset{\sim}{\gamma}_{tmt1} \cdots \underset{\sim}{\gamma}_{tmta}]'$ contains all of the parameters for the regression of the response variable on time. The ordinary least squares estimator for $\text{vec}(\underset{\sim}{\gamma}')$ is

$$\text{vec}(\hat{\underset{\sim}{\gamma}}') = [(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}' \otimes (\underset{\sim}{M}'\underset{\sim}{M})^{-1}\underset{\sim}{M}']\ \text{vec}(\underset{\sim}{y}').$$

For the repeated measures model with a treatments arranged in b randomized blocks, the covariance matrix of $\text{vec}(\underset{\sim}{y}')$ is block diagonal,

$$\text{Var}(\text{vec}(\underset{\sim}{y}')) = [(\underset{\sim}{I} \otimes \underset{\sim}{1}\underset{\sim}{1}') \otimes \underset{\sim}{\Sigma}_{\beta}] + \underset{\sim}{I} \otimes \underset{\sim}{\Sigma}_{\epsilon}, \text{ where } \underset{\sim}{1} = [1\ 1\ \ldots\ 1]' \text{ contains } a \text{ elements.}$$

The covariance matrix of $\text{vec}(\hat{\underset{\sim}{\gamma}}')$ is then

$$\text{Var}(\text{vec}(\hat{\underset{\sim}{\gamma}}')) = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'(\underset{\sim}{I} \otimes \underset{\sim}{1}\underset{\sim}{1}')\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-1} \otimes (\underset{\sim}{M}'\underset{\sim}{M})^{-1}\underset{\sim}{M}'\Sigma_{\beta}\underset{\sim}{M}(\underset{\sim}{M}'\underset{\sim}{M})^{-1}$$
$$+ (\underset{\sim}{X}'\underset{\sim}{X})^{-1} \otimes (\underset{\sim}{M}'\underset{\sim}{M})^{-1}\underset{\sim}{M}'\underset{\sim}{\Sigma}_{\epsilon}\underset{\sim}{M}(\underset{\sim}{M}'\underset{\sim}{M})^{-1}.$$

If the matrix of time coefficients, $\underset{\sim}{M}$, is orthogonal, then the covariance matrix reduces to

$$\text{Var}(\text{vec}(\hat{\gamma}')) = (X'X)^{-1}X'(I \otimes 1\,1')X(X'X)^{-1} \otimes M'\Sigma_\beta M + (X'X)^{-1} \otimes M'\Sigma_\epsilon M.$$

In the forest nutrition experiment we are interested in estimating the mean, linear, and quadratic terms for N=336 kg/ha and P=28 kg/ha. The linear combination of interest is $L\hat{\gamma}$ where

$$L = \frac{1}{12}[12 \;\; 3 \;\; 3 \;\; 3 \;\; 3 \;\; -1 \;\; 2 \;\; -1 \;\; -1 \;\; 2 \;\; -1 \;\; -1 \;\; 2 \;\; -1 \;\; 3 \;\; 6 \;\; 3].$$

The covariance matrix of such a linear combination of the coefficients is then

$$\text{Var}(L\hat{\gamma}) = M'\Sigma_\beta M \otimes L(X'X)^{-1}X'((I \otimes 1\,1')X(X'X)^{-1}L' + M'\Sigma_\epsilon M \otimes L(X'X)^{-1}L'.$$