Repeated measures experiments in forestry: focus on analysis of response curves

M. P. MEREDITH

Miami Valley Laboratories, The Procter and Gamble Company, P.O. Box 398707, Cincinnati, OH 45239, U.S.A.

AND

S. V. STEHMAN[1]

State University of New York, College of Environmental Science and Forestry, 211 Marshall Hall, Syracuse, NY 13210, U.S.A.

Received September 10, 1990

Accepted January 21, 1991

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.

Treatment effects over time are frequently investigated using repeated measures designs, but analyses of these experiments frequently fail to address a primary objective of collecting data over time, namely description of the response curve. The analysis advocated in this paper utilizes the intrinsic continuity of the repeated measures factor by focusing on response curves. Treatments are compared by analyzing estimated coefficients of response curves proposed by the investigator. This approach provides more information on treatment effects than analyses that compare treatments separately at each time period. Analysis of estimated coefficients is easier to interpret than multivariate analyses of variance and does not require often biologically implausible assumptions of split-plot analyses currently in vogue. An example describing effects of aluminum on sugar maple (*Acer saccharum* Marsh.) seedling growth illustrates the method.

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

Les effets des traitements dans le temps sont fréquemment évalués en utilisant des dispositifs à mesures répétées mais l'analyse de ces expériences manque souvent d'adresser un objectif premier de la récolte de données dans le temps, à savoir la description de la courbe des réponses. L'analyse présentée dans ce papier utilise la continuité intrinsèque du facteur mesures répétées en focalisant sur les courbes de réponse. Les traitements sont comparés en analysant les coefficients estimés des courbes de réponse proposées par l'investigateur. Cette approche apporte plus d'informations sur les effets des traitements que les analyses qui comparent les traitements séparément pour chaque période de temps. L'analyse des coefficients estimés est plus facile à interpréter que l'analyse multidimensionnelle de variance et ne requiert pas les hypothèses souvent non plausibles sur le plan biologique des analyses en tiroir en vogue présentement. Un exemple décrivant les effets de l'aluminium sur la croissance de semis d'érable à sucre (*Acer saccharum* Marsh.) illustre la méthode.

[Traduit par la rédaction]

Introduction

The variety of analyses and experiments labelled as repeated measures (RM) in textbooks and journals creates confusion for researchers trying to select an appropriate statistical strategy. Much of this well-documented and extensive RM literature is directed toward nonforestry applications in which the experimental protocols and objectives are much different from those common to forestry research. Not surprisingly, review of the forestry literature shows that statistical analyses of forestry RM data are frequently not as informative as they could be. Recent emphasis on various split-unit approaches to analysis of RM experiments (Moser et al. 1990; Crowder and Hand 1990) focuses on the error or correlation structure of RM data. This focus distracts from the primary objective of many forestry RM studies, which is to investigate treatment effects on the nature and behavior of response functions over time or space. We advocate an approach to statistical analysis of RM experiments that focuses on characteristics of these response curves.

RM designs in forestry typically consist of observations at fixed intervals of time (e.g., years or growing seasons) or space (e.g., soil depth) for plots or stands undergoing some treatment regimen. Our use of the term repeated measures corresponds to Koch et al.'s (1988) "longitudinal studies" classification or Finney's (1990) type I classification. Longitudinal studies are experiments in which the same experimental units (EUs) are measured at several times or locations, but the treatment applied to an EU does not change. We also require that the same variable be measured at each time period. These restrictions eliminate from consideration many RM experiments described for psychology, education, medical, dairy science, and behavior applications in which treatments are changed at successive time periods (e.g., crossover trials or change-over designs), or a set of tests, each measuring a different variable, is given to each subject (Johnson and Wichern 1988; Morrison 1989). We also restrict attention to experiments in which the RM factor is continuous.

A clear distinction between experimental units and sampling units is necessary for an appropriate analysis of RM data. Recall that an EU is the smallest unit of experimental material to which a treatment is randomized. In forestry experiments an EU may be an individual tree, a plot of trees, or some other grouping of sampling units. In the case where a plot of trees is an EU, the individual trees may be thought of as sampling units. Another frequently encountered situation is a subset of sampling units being taken destructively, often randomly, from an EU at each time. Sometimes these two techniques will be combined, as is the case when, for

[1]Author to whom all correspondence should be addressed.

Printed in Canada / Imprimé au Canada

TABLE 1. Statistical layout of RM data

		Time				
Treatment	EU	1	2	3	...	t
1	1	y_{111}	y_{112}	y_{113}	...	y_{11t}
1	2	y_{121}	y_{122}	y_{123}	...	y_{12t}
.	.	.	.	.		.
.	.	.	.	.	...	.
.	.	.	.	.		.
1	j	y_{1j1}	y_{1j2}	y_{1j3}	...	y_{1jt}
.	.	.	.	.		.
.	.	.	.	.	...	.
.	.	.	.	.		.
1	n	y_{1n1}	y_{1n2}	y_{1n3}	...	y_{1nt}
.	.	.	.	.		.
.	.	.	.	.		.
.	.	.	.	.		.
k	1	y_{k11}	y_{k12}	y_{k13}	...	y_{k1t}
k	2	y_{k21}	y_{k22}	y_{k23}	...	y_{k2t}
.	.	.	.	.		.
.	.	.	.	.	...	.
.	.	.	.	.		.
k	j	y_{kj1}	y_{kj2}	y_{kj3}	...	y_{kjt}
.	.	.	.	.		.
.	.	.	.	.	...	.
.	.	.	.	.		.
k	n	y_{kn1}	y_{kn2}	y_{kn3}	...	y_{knt}

example, diameter at breast height is measured on all trees within an EU, but several individual trees are removed for determining total biomass at each time. This paper is aimed at those forestry experiments where the same EUs are measured repeatedly through time without substantively affecting the EU at each measurement.

A distinguishing feature of RM designs is that levels of the RM factor cannot be assigned at random. For example, levels of time cannot be assigned randomly because the 3rd week of growth cannot be assigned to the 1st week of the experiment. This absence of randomization frequently results in observations that are serially correlated, and this invalidates the within-time mean square as a proper measure of experimental error. Choosing a method of analysis and corresponding estimates of appropriate error terms must be done with this serial correlation in mind.

We focus on time as the RM factor, but the results apply equally well to depth or other spatial factors. The almost universal practice of plotting treatment means versus time to describe results of RM experiments would argue that researchers are interested in the behavior of the response over time. The statistical analysis recommended herein reflects this interest in the response function. Repeated measures on each EU are replaced by a few estimated coefficients of a response curve, and these estimated coefficients are then used as data for comparison of treatments. A simple quadratic regression model, with the linear coefficient indicating trend and the quadratic coefficient describing the rate of increasing or decreasing behavior, provides one example of the type of response function useful for comparing treatments. The analysis strategy directly compares treatments using response curves selected by the researcher to satisfy the objectives of the experiment.

Data and model

For simplicity, assume that k treatments are allocated to EUs in a completely randomized design, and that the k treatments are equally replicated with n EUs per treatment. Measurements are taken at t equally spaced times for each of the EUs, resulting in a total of knt observations (Table 1).

For RM data, the observation y_{ijm} is modelled as the sum of an effect due to the treatment i at time m, μ_{im}, and an additive, homoscedastic, normally distributed random error component, ε_{ijm}:

$$[1] \quad y_{ijm} = \mu_{im} + \varepsilon_{ijm}$$

for $i = 1,2,\ldots,k$ treatments, $j = 1,2,\ldots,n$ replications of EUs within each treatment, and $m = 1,2,\ldots,t$ time periods. The response y_{ijm} may represent a transformation from the original scale of measurement in order to more closely meet these assumptions. Often it is helpful to assume some functional form for μ_{im}, such as $\mu_{im} = f(\beta_i,\mathrm{m})$, where, for example, $f(\beta_i,m)$ is a polynomial of degree less than t, and β_i is the parameter vector of coefficients of the polynomial. A quadratic response function would be modelled as

$$[2] \quad \mu_{im} = \beta_{i0} + \beta_{i1}m + \beta_{i2}m^2$$

Methods of analysis

Methods used to analyze RM experiments may be broadly grouped into (*i*) separate analyses at each time period; (*ii*) split-unit univariate analysis; (*iii*) multivariate analysis; and (*iv*) analysis of coefficients of functions estimated from the t repeated measures on each EU. Numerous other suggestions for analyzing RM studies exist, most of which are variations on the four themes just listed.

By far, the first method is most commonly used to analyze RM data in forestry. Of the 23 articles meeting our definition of a RM experiment appearing in *Forest Science* (1987) and the *Canadian Journal of Forest Research* (1988), 16 provided descriptive or inferential statistics for each time period separately or the last time period only, two papers reported split-unit analyses of RM experiments, and two employed multivariate analyses. Three papers reported analyses of coefficients of time response curves or other curves modelling growth response. Two recent review articles (Moser et al. 1990; Potvin et al. 1990) reported both split-unit and multivariate analyses. Other forestry RM studies cited for illustration in subsequent sections were found by a directed search of articles with titles hinting at use of an RM design.

Separate analyses at each time point

The simplest application of this approach is to plot the response means and standard errors at each time to display the observed response through time (see Fig. 1) and to indicate any trends in variance over time (Edmonds and McColl 1989; and Radwan et al. 1984). Performing a statistical analysis on the last time period only is also common (O'Reilly et al. 1989; Arnott and Macey 1985; Thornton et al. 1986). When a statistical analysis is performed within each time period instead of just at the last time period, analysis of variance (ANOVA) followed by a multiple comparisons procedure is frequently used (Mahoney et al. 1984; Kelliher et al. 1980; Barclay and Brix 1985). Qualitative descriptions concerning treatment responses over time are based on the t separate ANOVA's, but are usually not supported by statistical inferences.

The foremost criticism of employing separate ANOVA's at each time period to compare treatments is that serial correlation between observations over time is ignored, so the separate tests are not independent. Inferences about the response through time are made indirectly, and the shape of the response curve is inadequately

TABLE 2. ANOVA table for univariate split-unit analysis

Source	df	MS	F
Treatment	$k-1$	MS(treatment)	F_1 = MS(treatment)/MSE a
Error a	$k(n-1)$	MSE a	
Time	$t-1$	MS(time)	F_2 = MS(time)/MSE b
Time × treatment	$(k-1)(t-1)$	MS(treatment × time)	F_3 = MS(treatment × time)/MSE b
Error b	$k(n-1)(t-1)$	MSE b	

described. Sometimes the difference between the final and initial observations is analyzed to investigate the time effect, but if $t \geq 3$, this simply represents an inefficient way to estimate linear trend as it ignores the intermediate observations.

The separate analysis approach is also often used to locate the point in time at which treatment responses become significantly different. A problem with this procedure is that the time at which significant differences are detected is a function of the number of replications of EUs. A change in the number of replications changes the point in time at which "significance" is detected. Comparison of response curves provides a more effective analysis because it utilizes the continuity of the RM factor time. If response curves are significantly different, then treatment differences exist across time, not just at the discrete times at which measurements were taken. Treating time as a continuous or quantitative factor instead of analyzing time periods separately is analogous to a standard recommendation not to use multiple comparisons of means with quantitative treatment structures.

Univariate split-unit analysis

Split-unit (or split-plot) experiments are well known in forestry. We review the standard analysis to illustrate its application to RM data. The univariate split-unit approach treats the t repeated measures for each EU as t subunit measurements on each EU or whole unit (Winer 1971; Freund et al. 1986). If k treatments are applied to n EUs per treatment in a completely randomized design, F_1 (Table 2) tests for treatment main effects, while F_2 and F_3 test for time main effects and treatment by time interactions, respectively (Steel and Torrie 1980). Standard errors for various main and simple effect comparisons of the split-plot analysis are available in Cochran and Cox (1957, p. 298). Modifications of the basic table are straightforward for analysis of more complicated designs. Adams et al. (1987) and Moser et al. (1990) provide examples of split-unit analyses of RM data in forestry.

When applied to RM data, the split-unit approach has two major disadvantages. First, the tests for time effects and time by treatment interactions (Table 2) are often of less interest to researchers than the comparison and description of response curves. Testing of the time main effect is rarely of interest in forestry because we know *a priori* that changes occur through time, as in growth. Second, the within-time residual mean square, error b, necessary to obtain these comparisons and descriptions is generally inappropriate owing to serial correlation of observations over time (Hearne et al. 1983) and lack of randomization of the time factor.

This latter disadvantage is often remediated by requiring that an assumption of spherically symmetric errors is satisfied (Huynh and Feldt 1970). This sphericity assumption is for practical purposes equivalent to an assumption of equicorrelation or compound symmetry (Wallenstein 1982), which lacks a biologically plausible motivation. Observations close in time will have more highly correlated error components than observations farther apart in time (Finney 1990). Violation of compound symmetry due to the presence of typically strong, but even moderate, serial correlations in forestry RM data results in overly liberal tests (too many significant results) of the hypotheses associated with time and time by treatment interaction (Hearne et al. 1983; Huynh and Feldt 1976).

Considerable effort, particularly in psychometric applications, has been devoted to developing conservative and adjusted nominal-level approaches to split-unit analysis when the sphericity assumption is violated (Greenhouse and Geisser 1959; Huynh and Feldt 1970; Huynh and Feldt 1976; Rogan et al. 1979; Monlezun et al. 1984). The typical remedy is either to construct a conservative F-test (Box 1954*a*, 1954*b*; Wallenstein and Fleiss 1979) or to estimate the "approximate degrees of freedom" depending upon an estimate of a measure for departure from sphericity (Huynh and Feldt 1976). These alternatives are currently standard fare in the output of most major statistical packages (Freund et al. 1986; Dixon et al. 1988).

Another common strategy for the split-unit approach is to conduct preliminary tests to ascertain the validity of the sphericity and homogeneity assumptions. These methods usually have very low power for testing the validity of the sphericity assumption using Mauchley's (1940) criterion (Grieve 1984) and are also very sensitive to departures from multivariate normality. However, the level of the F-test for testing time and time by treatment effects is sensitive to extremely small departures from the sphericity assumption (Keselman et al. 1980; Boik 1981). Another important consequence of pretesting using Mauchley's test for sphericity is that significance levels reported later on in the analysis are biased due to this conditional inference (Bancroft and Han 1977). To avoid such problems, Boik (1981, p. 254) recommends that "A reasonable alternative is to employ a separate estimate of experimental error for each contrast among the repeated measures. The loss in degrees of freedom is more than compensated for by the gain in control over test size and power." This recommendation is precisely the idea promoted herein of analyzing estimated coefficients of the response function (i.e., contrasts) fitted to each EU.

One situation in which the split-unit approach may be used is when the subset of sampling units are taken destructively and at random from an EU at each time. It may be argued that the time factor can be randomized as a split-unit factor and the compound symmetry assumption will be valid, just as it is in a randomized complete block or split-plot experiment. See Potvin et al. (1990) for further discussion of this issue.

Multivariate analysis

The t repeated measures on each EU may be correctly regarded as a t-variate response vector and analyzed via multivariate methods provided t does not exceed the number of EUs for each treatment. Assumptions of normality and equal variance–covariance matrices across the k treatment populations are required for valid hypothesis testing, but an advantage of multivariate analysis is that no assumptions about the structure or pattern of the variance–covariance matrix are necessary. Population comparisons can proceed using either standard multivariate analysis of variance (MANOVA) (Milliken and Johnson 1984; Freund et al. 1986), profile analysis (Danford et al. 1960; Cole and Grizzle 1966; Davidson 1980; Morrison 1989), or multivariate analysis of growth curves (Rao 1958, 1965; Potthoff and Roy 1964; Grizzle and Allen 1969).

A major disadvantage of multivariate analysis is lack of power due to having to estimate the $t(t-1)/2$ parameters of the variance–covariance matrix. This becomes an important consideration when t is large and n is small, a situation common to long-term forestry

TABLE 3. Estimated coefficients of a quadratic polynomial response curve for each EU used as data for analysis of a RM experiment

Treatment	EU	Z	B	Q
1	1	z_{11}	b_{11}	q_{11}
1	2	z_{12}	b_{12}	q_{12}
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
1	j	z_{1j}	b_{1j}	q_{1j}
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
1	n	z_{1n}	b_{1n}	q_{1n}
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
k	1	z_{k1}	b_{k1}	q_{k1}
k	2	z_{k2}	b_{k2}	q_{k2}
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
k	j	z_{kj}	b_{kj}	q_{kj}
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
k	n	z_{kn}	b_{kn}	q_{kn}

trials. Further, multivariate statistical techniques are less well understood by researchers. This confusion sometimes leads researchers to analyses often intended to address objectives fundamentally different from objectives of RM experiments in forestry. For example, the hypotheses usually tested in multivariate profile analysis fail to exploit the underlying continuity of the response functions in time, and thus only indirectly answer questions of interest in forestry. Moreover, profiles frequently represent the outcome of different tests given to the same subject, which is not the experimental situation in most forestry RM studies. However, Kenward (1987) suggests a profile analysis approach applicable to continuous RM factors. Profile analysis may be a reasonable approach for discontinuous RM factors, such as soil horizons.

Analysis of estimated coefficients

The approach advocated for analyzing RM data focuses upon analyzing coefficients of functions fitted to the t repeated measures for each EU (Wishart 1938; Box 1950; Rowell and Walters 1976; Zerbe 1979; and Yates 1982). Restrictions on randomization (design features) and covariates create no special problems in this analysis.

It is important to note that each EU provides a set of estimated coefficients that are then treated as primary data. This set of coefficients generally results in a reduction in the dimensionality of the t-dimensional response vectors, which serves as an initial simplification to the problem. The investigator should choose a response function guided by the objectives of the experiment, so analyzing the estimated coefficients of this function will provide direct evidence for, or against, hypotheses initially set forth in the experiment.

The analysis illustrated in this paper will employ simple polynomial response functions to approximate the response curve. Use of a polynomial response does not imply acceptance of this function as the true underlying response curve. Rather, polynomials can be used to investigate characteristics of the response curve such as trend, humps and troughs, or asymmetry (Dawkins 1983). In the absence of prior knowledge about the form of the response curve, or when observing a restricted range of a more complicated response curve, a polynomial often will serve as a useful first approximation. However, if interest lies in curves that are S-shaped or have asymptotes within the time interval studied, then polynomials would provide a poor approximation.

Often experimental objectives will specify *a priori* a model for the response curve so that a simple polynomial approximation will not be necessary. If exponential decay or some other asymptotic behavior is expected, coefficients from an appropriate nonlinear model fitted to each EU may be used as data for the analysis. If the investigator is interested in time until a maximum response occurs, a quadratic model could be reparametrized in terms of time until maximum response (Meredith and Rubin 1990) and treatments compared using estimated coefficients of this model as data. Yates (1982) and Finney (1990) provide additional examples of other forms of response functions that could be used for RM data, and Potvin et al. (1990) suggest examples of response curves in physiological ecology.

The concepts behind the analysis of estimated coefficients are not unknown in forestry. Leak (1966) recognized the presence of correlation among observations within permanent remeasurement plots and described how to calculate the standard errors of slope estimates used to predict growth from plot variables. Bolstad and Allen (1987) and Buford and Burkhart (1987) compared treatments by analysis of regression coefficients for a model relating height to diameter. Based on objectives specified for their experiment, Overton and Ching (1978) identified several growth trajectory parameters characterizing growth response curves of Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco). Estimates of these parameters on each EU were analyzed to assess treatment differences. Several papers have investigated methods for estimating yield functions from permanent plot data, taking into account serial correlation over time (Ferguson and Leech 1978; Davis and West 1981; and Gregoire 1987). These methods are more concerned with efficient estimation of parameters than comparison of response functions, but the estimation procedures and models could be applied to comparative RM studies.

To illustrate the analysis of estimated coefficients method, suppose there are n replications of EUs within each of the k treatments. These EUs are observed at three equally spaced points in time. A quadratic polynomial will be used to model the response curve of the means over time. The comparison of treatment differences proceeds by estimating the coefficients of the quadratic polynomial for each EU. Orthogonal polynomial coefficients are preferred over the correlated partial regression coefficient estimates of the parametrization in eq. 2. With equally spaced observations in time, the linear component of the response function for the jth EU in the ith treatment is estimated by $b_{ij} = y_{ij3} - y_{ij1}$, and the quadratic component by $q_{ij} = y_{ij1} - 2y_{ij2} + y_{ij3}$. Another variable, $z_{ij} = (y_{ij1} + y_{ij2} + y_{ij3})/3$, is used to obtain the whole plot analysis, an analysis of the averages over time. The random variables Z, B, and Q become data for the analysis of coefficients (Table 3).

Serial correlation is of concern for measurements over time within the same EU, but observations or functions of

TABLE 4. ANOVA tables for analysis of estimated coefficients approach

Source	df	MS(Z)	F	MS(B)	F	MS(Q)	F
Mean	1	MSZ	F_{z1}	MS*B*	F_{b1}	MS*Q*	F_{q1}
Treatment	$k-1$	MSTZ	F_{z2}	MST*B*	F_{b2}	MST*Q*	F_{q2}
Error	$k(n-1)$	MSEZ		MSE*B*		MSE*Q*	

NOTE: *F*-statistics are obtained by dividing the effect mean square (MS) by the appropriate error mean square (MSE) for the coefficients representing the mean (Z), linear (B), and quadratic (Q) effects.

observations on different EUs are statistically independent by virtue of random assignment of treatments to EUs. Analysis of the estimated coefficients of the response function may proceed using univariate ANOVA, described subsequently, or MANOVA techniques (Capizzi and Burton 1978; Evans and Roberts 1979; Allen 1983; Allen et al. 1983).

Three ANOVA tables are calculated, one for each of Z, B, and Q (Table 4). Analysis of Z is equivalent to the usual whole-unit comparison of the treatment main effect. To analyze the effect of the RM factor time, the statistic F_{b2} tests the hypothesis that the linear component of the response function is the same for all k treatments. Rejection of the hypothesis of no treatment differences in the linear component of the time effect means that a statistically significant treatment by linear time interaction is present. That is, the linear trend in time, or slope, is not the same for all k treatments. If this interaction is not significant, then F_{b1} tests if the overall linear trend is statistically significant, or equivalently, if the coefficient of the linear component of time is zero when averaged over all k treatments.

The analysis of the quadratic component proceeds analogously using the variable Q. The statistic F_{q2} tests if a significant treatment by quadratic time interaction (different curvature of the response function) exists. If this interaction is not significant, F_{q1} provides a test determining if a common quadratic effect is zero. If both tests of the quadratic coefficient are not significant, the simpler linear model may be used.

Several key features of the analysis of coefficients need to be emphasized. First, the analysis directly addresses the shape of the response function by examining easily interpreted coefficients representing linear and quadratic trend of the response function. In analyzing variables B and Q, the treatment line of the ANOVA table represents an interaction effect, time by treatment, not a treatment main effect. Analysis of these variables partitions the degrees of freedom of the time by treatment interaction (Table 2) into linear time by treatment and quadratic time by treatment interactions, and MSE*B* and MSE*Q* (Table 4) are appropriate error terms for testing or estimating these interaction effects. Similar partitioning of the time by treatment interactions can be accomplished in split-unit analysis (Table 2), but without further assumptions on the error structure, this approach fails to correctly partition the error term. Analyzing orthogonal polynomial contrasts of means in the split-unit approach, as described by Moser et al. (1990), provides the same estimates of response curve parameters. But the error used for hypothesis testing or interval estimation in the split-unit approach requires the sphericity assumption or other modification, whereas the analysis of coefficients approach does not.

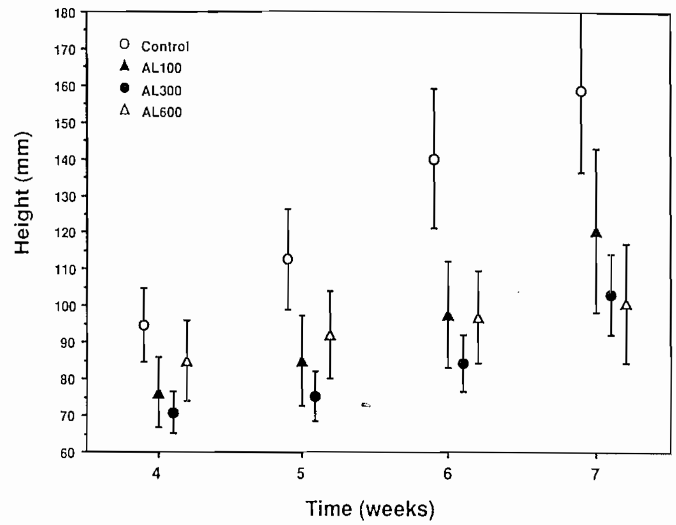

FIG. 1. Mean height of sugar maple seedlings over time for each concentration of aluminum (μM). Each mean is calculated from 17 observations. Standard error bars illustrate increasing variance over time.

Because the response functions (or any other combination of interest, linear or nonlinear) are specified by the investigator to answer questions about the process under study, the analysis of coefficients approach is not preprogrammed into any statistical software package. If these response functions can be approximated by orthogonal polynomial contrasts, then several statistical software packages provide the decomposition of sums of squares, but not estimates of response curve parameters. The usual tables of orthogonal polynomial coefficients cannot be used if observations are obtained at unequal time intervals. Robson (1959) describes how these coefficients may be estimated for unequal spacing.

If the estimated coefficients analyzed are not independent, analysis options include MANOVA or separate univariate analyses of the coefficients with adjustment of the error levels using Bonferroni bounds (Feller 1968). Meredith (1988) describes a biologically plausible specification of the variance–covariance matrix for which even- and odd-order orthogonal polynomial coefficients are independent, and no adjustments are necessary. Note that the multivariate analysis of estimated coefficients is easier to conduct and interpret than the usual MANOVA of the original t-variate response.

We have not addressed the problem of incomplete data on some EUs. If only a small percentage of EUs have incomplete records, the analysis could be performed on only those EUs with the necessary data to estimate the response function of interest. Care should be taken so that selection bias is not introduced by this approach (Koch et al. 1988). Berk (1987) and the references contained therein provide a starting point for reading about more complicated analyses for incomplete data.

Example: effect of aluminum on growth of sugar maple

The analysis and interpretation of an example illustrates the analysis of coefficients approach to RM data. Thornton et al. (1986) studied the effect of aluminum on growth of sugar maple (*Acer saccharum* Marsh.) seedlings. A subset of data from one of their experiments is used here. Height

TABLE 5. Statistical analysis of estimated coefficients of a cubic polynomial response curve for sugar maple height data

(A) Whole-unit analysis (Z)

Source	df	MS	p-value
Mean	1	673 913	<0.01
Aluminum	3	5 923	0.11
Residual	64	2 806	

(B) Analysis of RM factor

		Contrast					
		Linear (B)		Quadratic (Q)		Cubic (C)	
Source	df	MS	p-value	MS	p-value	MS	p-value
Mean	1	11 622.4	<0.01	177.9	0.14	0.23	0.75
Contrast × aluminum	3	846.5	0.03	90.4	0.33	5.56	0.08
Residual	64	275.9		78.8		2.40	

measurements (mm) from weeks 4 through 7 were obtained from seventeen 23-day-old seedlings grown in solution at four aluminum concentrations, 0, 100, 300, and 600 μM. The orthogonal polynomial coefficients used to calculate the linear (B), quadratic (Q), and cubic (C) contrasts are:

[3] $b_{ij} = (-3y_{ij4} - y_{ij5} + y_{ij6} + 3y_{ij7})/10$

[4] $q_{ij} = (y_{ij4} - y_{ij5} - y_{ij6} + y_{ij7})/4$

[5] $c_{ij} = (-y_{ij4} + 3y_{ij5} - 3y_{ij6} + y_{ij7})/6$

where the divisors are used to obtain coefficients as change on a per week basis.

Figure 1 displays the treatment means (and standard errors) at each time. As expected for most growth data, the standard errors of the means increase over time. Even if variance heterogeneity or serial correlation exist between time points in the original data, the orthogonal polynomial coefficients remain unbiased. More complicated estimation procedures may be used to obtain more efficient estimates of the coefficients. A logarithmic transformation would be well motivated for these data because growth data often follow a lognormal distribution (Aitchison and Brown 1957). However, residual plots from the analysis of the estimated coefficients showed no major departure from variance homogeneity for the time interval studied, so the analysis was conducted on the original scale for ease of analysis and interpretation.

ANOVA tables for the whole unit analysis of means across weeks (Z), and variables B, Q, and C are presented in Table 5 (the SAS program and data are provided in the Appendix). The nonsignificant main effect of aluminum ($p = 0.11$) must be interpreted cautiously because the linear time by treatment interaction is significant ($p = 0.03$), indicating a difference (lack of parallelism) in the linear component of the response curves for different aluminum concentrations. Neither the quadratic time by treatment interaction ($p = 0.33$) nor quadratic time main effect is significant ($p = 0.14$). Note that a plausible biological interpretation of the cubic time by aluminum interaction ($p = 0.08$) is difficult given the early phase of growth investigated in this experiment. This result may be due to different technicians taking measurements at different times, correlation between linear and cubic coefficients (Meredith 1988), or random chance.

Given the preceding results, a separate linear response curve is deemed adequate. The prediction equation for mean cumulative height (mm) of the ith treatment is given in terms of sequential regression coefficients:

[6] $\hat{y}_i = \bar{z}_i + \bar{b}_i(\text{week} - 5.5)$

where i = 0, 100, 300, or 600 μM, $\bar{z}_i$ and $\bar{b}_i$ are treatment means of the variables Z and B for each aluminum concentration (Table 6), and 5.5 is the mean of the variable week. Standard errors of $\bar{z}_i$ and $\bar{b}_i$ are obtained from the respective square roots of MSEZ and MSEB in Table 5. For illustration, the prediction equation for mean cumulative height at i = 100 μM is (standard errors in parentheses below the estimates):

[7] $$\underset{}{\hat{y}_{100}} = \underset{(12.8)}{94.8} + \underset{(4.0)}{14.5}(\text{week} - 5.5)$$

The estimated coefficients, $\bar{z}_i$ and $\bar{b}_i$, are exactly those obtained by ordinary least squares regression of cumulative growth against time for each aluminum concentration. However, standard errors obtained via this regression would be incorrect due to serial correlation of observations.

The levels of aluminum represent a quantitative treatment factor, so further analysis of treatment effects can proceed by investigating the dose–response to aluminum. The 3 df linear time by aluminum interaction can be partitioned by standard contrast methodology (section 12.8, Snedecor and Cochran 1980). The means used in these contrasts are means of estimated slopes for each treatment. The contrast testing for linear trend in the mean slopes is significant ($p < 0.01$), with no evidence of significant lack of fit from a linear trend ($p = 0.63$). Therefore, the positive linear trend in growth over time decreases linearly with increasing aluminum concentrations.

The linear prediction equation for the mean estimated slope, $\hat{\bar{b}}_i$, at each level of aluminum is given by

[8] $$\hat{\bar{b}}_i = \underset{(2.0)}{13.07} - \underset{(0.009)}{0.0253}(i - 250)$$

TABLE 6. Estimated means and standard errors for the treatment response averaged over time (Z) and the linear coefficient of the response curve (B)

Aluminum concn. (μM)	Contrast	
	Mean (Z)	Linear (B)
0	126.5	22.0
100	94.8	14.5
300	83.3	10.6
600	93.6	5.3
SE	12.8	4.0

where i is aluminum concentration and 250 is the mean aluminum concentration of the four treatments. Substituting $\hat{\bar{b}}_i$ for $\bar{b}_i$ in the prediction equation for height [6] leads to the formula

$$[9] \quad \hat{y}_i = \bar{y}_i + \underset{(2.0)}{[13.07} - \underset{(0.009)}{0.0253}(i - 250)]\,(\text{week} - 5.5),$$

where $\bar{y}_i$ is the mean height of seedlings at aluminum concentration i. By fixing either a value of week or aluminum, a simple linear regression results. For example, if week = 5, then:

$$[10] \quad \hat{y}_i = [\underset{(12.8)}{\bar{y}_i} - \underset{(1.0)}{6.53}] + \underset{(0.004)}{0.0126}\,(i - 250)$$

If the aluminum concentration is fixed, eq. 6 becomes

$$[11] \quad \hat{y}_i = \bar{z}_i + \hat{\bar{b}}_i\,(\text{week} - 5.5)$$

For example, if the aluminum concentration is 300, then

$$[12] \quad \hat{y} = \underset{(12.8)}{83.3} + \underset{(4.25)}{11.8}\,(\text{week} - 5.5)$$

The treatment response curves predicted by [11] are displayed in Fig. 2.

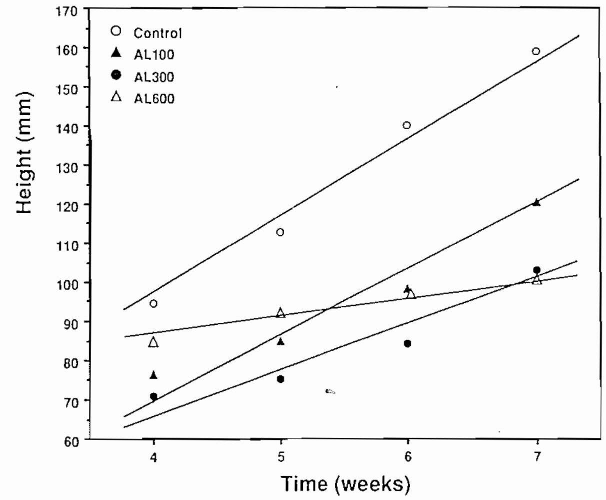

FIG. 2. Estimated response curves for cumulative height of sugar maple seedlings grown at four aluminum concentrations (μM) as predicted by eq. 11.

Conclusions

Researchers employing RM designs in forestry are most often interested in eliciting pertinent information about the behavior of the underlying response curves of the process under investigation. Analysis of coefficients of response curves provides this information directly. In general, we recommend analysis of coefficients of response curves hypothesized *a priori* to satisfy research objectives of the experiment. Focusing on response curves forces the researcher to clearly define the objective of investigating a treatment response over time or space. Even if no *a priori* curve is postulated, analysis of coefficients of an approximate response curve, such as a low-order polynomial, is a significant improvement over many currently reported RM analyses that provide no statistical description of the response curve, but rather report separate analyses at each time period.

Effective analysis of forestry RM experiments is sometimes hampered by adherence to MANOVA and split-unit analyses available in, and thus inadvertently promoted by, major statistical software packages and statistical textbooks written for nonforestry disciplines. We generally prefer the analysis of estimated coefficients to MANOVA of the t repeated measures on each EU because MANOVA often has low power, does not directly examine the form of the continuous response curve, and is more difficult to interpret than univariate analyses. The analysis of coefficients approach also does not require the difficult to verify or biologically suspect assumptions inherent in split-unit analyses. We concur with Koch et al.'s (1988) recommendation that split-unit analysis be used only "when sphericity can be presumed on the basis of subject matter knowledge and research design structure (e.g., split-unit experiments); otherwise, univariate analysis of within-subject functions is preferable."

Implicit in an analysis of coefficients approach is the mandate for simplicity in data analysis while directly addressing research objectives. Analysis of coefficients of response functions selected by the investigator to satisfy research objectives satisfies the simplicity mandate for the statistical user, while simultaneously providing an interpretable and powerful analysis of RM experiments.

Acknowledgments

The authors are grateful to Dr. Michail Schaedle for providing data used in the example, State University of New York, College of Environmental Science and Forestry, for travel support, and Procter and Gamble for continued support of research. C. A. Nowak, D. A. Maguire, and W. T. Federer provided valuable reviews of earlier versions of the manuscript. D. M. Umbach's comments on the last version and examples of response functions were especially helpful.

ADAMS, M.B., CAMPBELL, R.G., ALLEN, H.L., and DAVEY, C.B. 1987. Root and foliar nutrient concentrations in loblolly pine: effects of season, site, and fertilization. For. Sci. **33**: 984–996.

AITCHISON, J., and BROWN, J.A.C. 1957. The lognormal distribution with special reference to its uses in economics. Cambridge University Press, London.

ALLEN, O.B. 1983. A guide to the analysis of growth curve data with special reference to SAS. Comput. Biomed. Res. **16**: 101–115.

ALLEN, O.B., BURTON, J.H., and HOLT, J.D. 1983. Analysis of repeated measurements from animal experiments using polynomial regression. J. Anim. Sci. **57**: 765–770.

ARNOTT, J.T., and MACEY, D.E. 1985. Effect of supplemental light intensity on white spruce, Engelmann spruce, and mountain hemlock seedlings grown under an extended photoperiod. Can. J. For. Res. **15**: 295–300.

BANCROFT, T.A., and HAN, C.-P. 1977. Inference based on conditional specification: a note and a bibliography. Int. Stat. Rev. **45**: 117–127.

BARCLAY, H.J., and BRIX, H. 1985. Effects of high levels of fertilization with urea on growth of thinned and unthinned Douglas-fir stands. Can. J. For. Res. **15**: 730–733.

BERK, K. 1987. Computing for incomplete repeated measures. Biometrics, **43**: 385–398.

BOIK, R.J. 1981. A priori tests in repeated measures designs: effects of nonsphericity. Psychometrika, **46**: 241–255.

BOLSTAD, P.V., and ALLEN, H.L. 1987. Height and diameter growth response in loblolly pine stands following fertilization. For. Sci. **33**: 644–653.

BOX, G.E.P. 1950. Problems in the analysis of growth and wear curves. Biometrics, **6**: 362–389.

______ 1954*a*. Some theorems on quadratic forms in the study of analysis of variance problems. I. Effect of inequality of variance in the one-way classification. Ann. Math. Stat. **25**: 290–302.

______ 1954*b*. Some theorems on quadratic forms in the study of analysis of variance problems. II. Effect of inequality of variances and correlation between errors in the two-way classification. Ann. Math. Stat. **25**: 484–498.

BUFORD, M., and BURKHART, H.E. 1987. Genetic improvement effects on growth and yield of loblolly pine plantations. For. Sci. **33**: 707–724.

CAPIZZI, T., and BURTON, D.T. 1978. A statistical technique for analyzing time-response curves. *In* Biological data in water pollution assessment: quantitative and statistical analyses. *Edited by* K.L. Dickson, J. Cairns, Jr., and R.J. Livingston. American Society for Testing and Materials, Philadelphia, PA. pp. 137–149.

COCHRAN, W.G., and COX, G.M. 1957. Experimental designs. John Wiley & Sons, New York.

COLE, J.W.L., and GRIZZLE, J.E. 1966. Applications of multivariate analysis of variance to repeated measurements experiments. Biometrics, **22**: 810–828.

CROWDER, M.J., and HAND, D.J. 1990. Analysis of repeated measures. Chapman and Hall, New York.

DANFORD, M.B., HUGHES, H.M., and MCNEE, R.C. 1960. On the analysis of repeated measurements experiments. Biometrics, **16**: 547–565.

DAVIDSON, M. 1980. The multivariate approach to repeated measures. BMDP Statistical Software, Los Angeles, CA. Tech. Rep. No. 75.

DAVIS, A.W., and WEST, P.W. 1981. Remarks on "Generalized least squares estimation of yield functions" by I.S. Ferguson and J.W. Leech. For. Sci. **27**: 233–239.

DAWKINS, H.C. 1983. Multiple comparisons misused: why so frequently in response-curve studies? Biometrics, **39**: 789–790.

DIXON, W.J., BROWN, M.B., ENGELMAN, L., HILL, M.A., and JENNRICH, R.I. 1988. BMDP statistical software manual. University of California Press, Los Angeles.

EDMONDS, R.L., and MCCOLL, J.G. 1989. Effects of forest management on soil nitrogen in *Pinus radiata* stands in the Australian Capital Territory. For. Ecol. Manage. **29**: 199–212.

EVANS, J.C., and ROBERTS, E.A. 1979. Analysis of sequential observations with applications to experiments on grazing animals and perennial plants. Biometrics, **35**: 687–693.

FELLER, W. 1968. An introduction to probability theory and its applications. 3rd ed. Vol. 1. John Wiley & Sons, New York.

FERGUSON, I.S., and LEECH, J.W. 1978. Generalized least squares estimation of yield functions. For. Sci. **24**: 27–42.

FINNEY, D.J. 1990. Repeated measurements: what is measured and what repeats? Stat. Med. **9**: 639–644.

FREUND, R.J., LITTELL, R.C., and SPECTOR, P.C. 1986. SAS system for linear models. SAS Institute Inc., Cary, NC.

GREENHOUSE, S.W., and GEISSER, S. 1959. On methods in the analysis of profile data. Psychometrika, **24**: 95–112.

GREGOIRE, T.G. 1987. Generalized error structure for forestry yield models. For. Sci. **33**: 423–444.

GRIEVE, A.P. 1984. Tests of sphericity of normal distributions and the analysis of repeated measures designs. Psychometrika, **49**: 257–267.

GRIZZLE, J.E., and ALLEN, D.M. 1969. Analysis of growth and dose response curves. Biometrics, **25**: 357–381.

HEARNE, E.M., CLARK, G.M., and HATCH, J.P. 1983. A test for serial correlation in univariate repeated measures analysis. Biometrics, **39**: 237–243.

HUYNH, H., and FELDT, L.S. 1970. Conditions under which mean square ratios in repeated measurements designs have exact *F*-distributions. J. Am. Stat. Assoc. **65**: 1582–1589.

______ 1976. Estimation of the Box correction for degrees of freedom from sample data in the randomized block and split-plot designs. J. Educ. Stat. **1**: 69–82.

JOHNSON, R.A., and WICHERN, D.W. 1988. Applied multivariate statistical analysis. 2nd ed. Prentice-Hall, Englewood Cliffs, NJ.

KELLIHER, F.M., KIRKHAM, M.B., and TURNER, C.G. 1980. Stomatal resistance, transpiration, and growth of drought-stressed eastern cottonwood. Can. J. For. Res. **10**: 447–451.

KENWARD, M.G. 1987. A method for comparing profiles of repeated measurements. Appl. Stat. **36**: 296–308.

KESELMAN, H.J., ROGAN, J.C., MENDOZA, J.L., and BREEN, L.J. 1980. Testing the validity conditions of repeated measures *F* tests. Psychol. Bull. **87**: 479–481.

KOCH, G.G., ELASHOFF, J.D., and AMARA, I.A. 1988. Repeated measurements—design and analysis. *In* Encyclopedia of statistical sciences. Vol. 8. *Edited by* S. Kotz, N.L. Johnson, and C.B. Read. John Wiley & Sons, New York. pp. 46–73.

LEAK, W. 1966. Analysis of multiple systematic remeasurements. For. Sci. **12**: 69–73.

MAHONEY, M.J., SKELLY, J.M., CHEVONE, B.I., and MOORE, L.D. 1984. Response of yellow poplar (*Liriodendron tulipfera* L.) seedling shoot growth to low concentrations of O_3, SO_2, and NO_2. Can. J. For. Res. **14**: 150–153.

MAUCHLEY, J.W. 1940. Significance test for sphericity of a normal *n*-variate distribution. Ann. Math. Stat. **11**: 204–209.

MEREDITH, M.P. 1988. Results and remarks on analyzing biological repeated measures. Biometrics Unit, Cornell University, 337 Warren Hall, Ithaca, NY 14853, U.S.A. Tech. Rep. BU-956-M.

MEREDITH, M.P., and RUBIN, G. 1990. Some useful reparametrizations: linear to nonlinear models. Biometrics Unit, Cornell University, 337 Warren Hall, Ithaca, NY 14853, U.S.A. Tech. Rep. BU-1070-M.

MILLIKEN, G.A., and JOHNSON, D.E. 1984. Analysis of messy data. Lifetime Learning Publications, Belmont, CA.

MONLEZÚN, C.J., BLOUIN, D.C., and MALONE, L.C. 1984. Contrasting split plot and repeated measures experiments and analyses. Am. Stat. **38**: 21–31.

MORRISON, D.F. 1989. Multivariate statistical methods. 3rd ed. McGraw-Hill, New York.

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

O'REILLY, C., ARNOTT, J.T., and OWENS, J.N. 1989. Effects of photoperiod and moisture availability on shoot growth, seedling morphology, and cuticle and epicuticular wax features of container-grown western hemlock seedlings. Can. J. For. Res. **19**: 122–131.

OVERTON, W.S., and CHING, K.K. 1978. Analysis of differences in height growth among populations in a nursery selection study of Douglas-fir. For. Sci. **24**: 497–509.

POTTHOFF, R.F., and ROY, S.N. 1964. A generalized multivariate analysis of variance model useful especially for growth curve problems. Biometrika, **51**: 313–326.

POTVIN, C., LECHOWICZ, M.J., and TARDIF, S. 1990. The statistical analysis of ecophysiological response curves obtained from experiments involving repeated measures. Ecology, **71**: 1389–1400.

RADWAN, M.A., DEBELL, D.S., WEBSTER, S.R., and GESSEL, S.P. 1984. Different nitrogen sources for fertilizing western hemlock in western Washington. Can. J. For. Res. **14**: 155–162.
RAO, C.R. 1958. Some statistical methods for comparison of growth curves. Biometrics, **14**: 1–17.
______ 1965. Theory of least squares when the parameters are stochastic and its application to the analysis of growth curves. Biometrika, **52**: 447–458.
ROBSON, D.S. 1959. A simple method for constructing orthogonal polynomials when the independent variable is unequally spaced. Biometrics, **15**: 187–191.
ROGAN, J.C., KESELMAN, H.J., and MENDOZA, J.L. 1979. Analysis of repeated measurements. B. J. Math. Stat. Psychol. **32**: 269–286.
ROWELL, J.G., and WALTERS, D.E. 1976. Analyzing data with repeated observations on each experimental unit. J. Agric. Sci. **87**: 423–432.
SNEDECOR, G.W., and COCHRAN, W.G. 1980. Statistical methods. 7th ed. Iowa State University Press, Ames.
STEEL, R.G.D., and TORRIE, J.H. 1980. Principles and procedures of statistics: a Biometrical approach. 2nd ed. McGraw-Hill, New York.
THORNTON, F.C., SCHAEDLE, M., and RAYNAL, D.J. 1986. Effects of aluminum on the growth of sugar maple in solution culture. Can. J. For. Res. **16**: 892–896.
WALLENSTEIN, S. 1982. Regression models for repeated measurements: reader reaction. Biometrics, **38**: 849–850.
WALLENSTEIN, S., and FLEISS, J.L. 1979. Repeated measurements analysis of variance when the correlations have a certain pattern. Psychometrika, **44**: 229–233.
WINER, B.J. 1971. Statistical principles in experimental design. 2nd ed. McGraw-Hill, New York.
WISHART, J. 1938. Growth-rate determinations in nutrition studies with the bacon pig, and their analysis. Biometrika, **30**: 16–28.
YATES, F. 1982. Regression models for repeated measurements: reader reaction. Biometrics, **38**: 850–853.
ZERBE, G.O. 1979. Randomization analysis of the completely randomized design extended to growth and response curves. J. Am. Stat. Assoc. **74**: 215–221.

Appendix: Data and SAS program for analyzing sugar maple example

```
data maple;
  title 'Repeated measures analysis of maple data';
  input al t4 t5 t6 t7;
  mean = 1;
  z = (t4 + t5 + t6 + t7)/4;
  b = (-3 * t4 - t5 + t6 + 3 * t7)/10;
  q = (t4 - t5 - t6 + t7)/4;
  c = 3 * (-t4 + 3 * t5 - 3 * t6 + t7)/6;
  cards; /* Columns 2-5 have height (mm) at weeks 4-7 */
    0    60   62   78  104
    0    41   50   60   60
    0    85   97  115  120
    0    88   87   90   80
    0    66   65   80   95
    0   106  100  133  172
    0    61   65   65   54
    0    52   62   65   50
    0   194  210  250  300
    0    86   90  105  163
    0   140  183  283  310
    0    41   45   83  135
    0    99  190  280  262
    0   144  208  220  221
    0   117  155  208  323
    0   102  100  105  100
  100    70   86  100   90
  100   167  190  230  320
  100   116  120  157  291
  100    85  100  125  180
  100    90   85   90  150
  100    44   45   70   82
  100    71   68   70   65
  100    40   35   37   34
  100    29   30   32   32
  100    49   50   60   65
  100   116  125  130  115
  100   135  195  225  265
  100    73   80   85  117
  100    39   40   40   40
  100    21   25   32   30
  100    69   80   80   86
  100    84   90   95   85
  300    67   65   65   62
  300   119  145  167  234
  300    72   70   80  115
  300    81   80   83   85
  300    31   35   43   95
  300    50   48   55   81
  300    62   75  105  177
  300    50   47   50   62
  300   111  120  130  140
  300    68   68   72   62
  300    80   78   83   74
  300    72   90  105  115
  300    45   40   45  103
  300    67   80   85   82
  300   105  105  108  100
  300    50   65   75   85
  300    74   70   80   80
  600    76   95   95   92
  600   111  115  115   95
  600   190  200  225  300
  600   110  135  140  123
  600   105  135  143  142
  600    67   60   60   55
  600    51   45   55   65
  600    64   53   62   65
  600    43   43   45   30
  600    34   65   60   80
  600    68   95  100  120
  600    55   60   62   55
  600    42   38   40   45
  600   126  130  127  110
  600    72   68   75   70
  600   175  175  183  210
  600    52   52   60   54
  ;
/* ANOVA of estimated coefficients */
proc glm; class al;
  model z b q c = mean al / noint;
  lsmean al / stderr;
  /* Contrasts of estimated coefficients */
  estimate 'Average over Al'
                          mean 1 al 0.25 0.25 0.25 0.25;
  estimate 'Slope of contrast'
                  al -0.25 -0.15 0.05 0.35 / divisor = 210;
```