A GENERAL APPROACH FOR THE ANALYSIS OF REPEATED MEASURES EXPERIMENTS

A. P. VERBYLA and B. R. CULLIS

1. INTRODUCTION

Over the past decade, there has been great interest in developing methods for analysis of repeated measures data. The major papers in the area promote statistical modelling to extract structure from the data which is often of prime interest to the researcher. The papers by LAIRD and WARE [9], DIGGLE [5], and CULLIS and MCGILCHRIST [3], as well as those by the authors are but a few examples. The reason for these developments is the desire to provide relevant, comprehensive and comprehensible analyses for complex situations.

Our approach outlined below can be found in three papers, VERBYLA and CULLIS [15], CULLIS and VERBYLA [4] and VERBYLA and CULLIS [16]. We feel the approach is comprehensive because it handles

- complete or *incomplete* data
- time-dependent covariates
- between experimental unit dependence through *blocking or spatial correlation* as well as temporal correlation
- modelling of treatment contrasts using linear and nonlinear models and smoothing techniques

Our approach is relevant and comprehensible because we strive to elicit as much structure as possible from the data and hence to answer the questions of interest to the researcher. Our analyses in section 5 illustrate the value of the approach.

We carry out our analysis as follows.

- Treatment effects are examined at each time or a saturated linear model (full treatment structure) is fitted assuming independence.
- Using the residuals from the preliminary fitting, we use one or more of the diagnostics
  - Residual sum of squares and products/correlation matrix
  - correlogram
  - empirical semi-variogram (DIGGLE, [5])

  to determine a reasonable covariance structure.

- We use REML, Residual Maximum Likelihood to estimate the parameters in the mean and covariance structure. Diagnostics may again be used to check the fitted model.
- Treatment contrasts are modelled if appropriate.
- Tests of hypotheses are carried out as required.

All the computation is carried out using a MATLAB program written by the second author. **S** and **Splus** versions will soon be available, a FORTRAN version is available and a GENSTAT implementation is also likely in the near future.

2. GENERAL MODEL

Suppose we have complete data for n plots over p time points. We suppose the $n \times p$ data matrix Y has expectation

$$\mathrm{E}(Y) = DT$$

where D is a known $n \times r$ matrix which specifies the treatment design, and T is a matrix of unknown parameters, the treatment effects over time. In vector notation, using the vec() operator which stacks the columns of the matrix argument, we have

$$\mathrm{E}(y) = \mathrm{E}\{\mathrm{vec}(Y)\} = (I_p \otimes D)\mathrm{vec}(T) = X\tau. \tag{1}$$

We denote vector versions of matrices by lower case equivalents. We assume that

$$\mathrm{var}(y) = \sigma^2(\Omega_2 \otimes \Omega_1) = \sigma^2 H \tag{2}$$

so that the error structure is separable, as used in spatial analysis by MARTIN [10] and CULLIS and GLEESON [2]. This structure is also appropriate for variance components models or multistratum designs under certain assumptions concerning the block or random effects. The temporal correlation structure is specified by Ω_2, while the spatial or between unit correlation structure is specified by Ω_1. Restrictions must be placed on Ω_1 and Ω_2 to ensure identifiability of parameters in the covariance structure. Both Ω_2 and Ω_1 will usually be modelled parsimoniously; we denote the parameter vectors for Ω_2 and Ω_1 by γ_2 and γ_1 respectively, and let $\gamma = [\gamma_1' \ \gamma_2']'$.

Missing data are a common problem in many experiments, as observations are lost, not taken or experimental units inadvertently drop out. In

designed experiments it is often of interest to estimate the missing values. We follow HOUTMAN and SPEED [8] and write

$$Y = Y_1 + Y_2 \qquad \text{or} \qquad y = y_1 + y_2$$

where Y_1 and y_1 are the observed data with zeros in the positions where data were not observed and Y_2 and y_2 represent the unobserved data with zeros in the positions where data are observed. We estimate y_2 as well as the unknown parameters τ and γ.

Suppose we have m_i missing data at the ith time point and let $m = m_1 + \ldots + m_p$. For $j = 1, 2, \ldots, m_i$, let F_{ij} be a $n \times 1$ vector consisting of $(n-1)$ zeros and a 'one' for the experimental unit with the jth missing value at the ith time point. Let $F_i = [F_{i1} \ldots F_{im_i}]$ and ϕ_i be parameters defined by $Y_2 = [F_1\phi_1 \ldots F_p\phi_p]$ or

$$y_2 = \begin{bmatrix} F_1 & 0 & \ldots & 0 \\ 0 & F_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & F_p \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{bmatrix} = F\phi. \tag{3}$$

We incorporate the parameters ϕ into the linear model (1) by

$$Y_1 = DT - [F_1\phi_1 \; F_2\phi_2 \ldots F_p\phi_p] + E \quad \text{or} \quad y_1 = X\tau - F\phi + e \tag{4}$$

and estimate τ and ϕ, thereby estimating the missing values. If any $m_i = 0$, both the column size of F and the dimension of ϕ are reduced accordingly.

In matrix form, apart from the minus sign, the incomplete data adjustment to the linear model (4) corresponds to time-dependent covariates as discussed by VERBYLA [12] and CULLIS and VERBYLA [4]. In that case, the matrices F_i contain the values of the time-dependent covariates for the ith time point.

We can avoid the unknown or missing data by simply transforming y or y_1 to a vector which contains only the observed data. Let K be an $np \times (np - m)$ matrix which indicates which of the np possible observations were observed. Each column has a single unit entry, the rest being zeros, $K'F = 0$ and the matrix $[KF]$ is a permuted identity matrix. We need only consider $K'y$ for estimation and this is the approach adopted in VERBYLA and CULLIS [15]. Then

$$\mathrm{E}(K'y) = K'X\tau, \quad \text{and} \quad \mathrm{var}(K'y) = \sigma^2 K'HK. \tag{5}$$

If only a few observations are missing or estimation of missing data is desirable, it may be preferable to use (4). If there is a good deal of missing data and estimation of the missing data is not of importance or is clearly questionable, (5) may be preferred.

3. REML ESTIMATION

A discussion of REML estimation can be found in HARVILLE [7], COOPER and THOMPSON [1], CULLIS and MCGILCHRIST [3] and VERBYLA and CULLIS [15]; see also VERBYLA [13, 14]. Put simply, we use a two stage procedure. In the first stage we assume the variance parameters are known and estimate the mean parameters. At the second stage we use a restricted likelihood based on error contrasts to estimate the variance parameters as introduced by PATTERSON and THOMPSON [11]. The mean parameter estimates are then found by substituting the estimated variance parameters into the results of stage one.

Estimation of τ and ϕ for given H reduces to the generalised least squares criterion, and maximum likelihood for given H, to minimise

$$\begin{aligned}(y - X\tau)'H^{-1}(y - X\tau) &= (y_1 + y_2 - X\tau)'H^{-1}(y_1 + y_2 - X\tau) \\ &= (y_1 + F\phi - X\tau)'H^{-1}(y_1 + F\phi - X\tau).\end{aligned}$$

An iterative procedure is required for estimation of σ^2 and γ and Fisher's method of scoring is one approach for solving the REML equations.

4. RANDOM EFFECTS

Suppose the experimental units are grouped into b blocks. For complete data we have

$$Y = DT + AB + E$$

where as above D is the treatment design matrix and T is the matrix of unknown parameters corresponding to the treatment effects, A is an $n \times b$ indicator matrix of zeros and ones which defines the block structure for the experiment, and $\text{vec}(E)$ and $\beta = \text{vec}(B)$ have independent multivariate normal distributions with zero mean vector and covariance matrices $\sigma^2(\Omega_2 \otimes I_n)$ and $\sigma^2\lambda(\Omega_2 \otimes I_b) = \sigma^2\Gamma$, say, respectively. The assumption is that the block effects have covariance structure proportional to that of the errors in a similar manner to that proposed in the univariate case. Clearly the block effects over time will be connected and our assumption leads to a separable covariance matrix. In vector form if $Z = I_p \otimes A$ we can specify the mixed model as

$$\text{E}(y|\beta) = X\tau + Z\beta, \qquad \text{var}(y|\beta) = \sigma^2(\Omega_2 \otimes I_n) = \sigma^2 V$$

and so the unconditional distribution of y is normal with

$$\mathrm{E}(y) = X\tau, \qquad \mathrm{var}(y) = \sigma^2(\Omega_2 \otimes \Omega_1),$$

as in Section 2, but we have a special form for Ω_1, namely $\Omega_1 = I_n + \lambda AA'$.

There is no problem with incomplete data, as we again write $y = y_1 + F\phi$ and include ϕ in the estimation procedure.

The results can be extended to the case of several blocking factors by replacing AB by $A_1B_1 + \ldots A_kB_k$ where we assume $\mathrm{var}(\mathrm{vec}(B_i)) = \sigma^2\lambda_i(\Omega_2 \otimes I_{b_i})$ and the B_i are independent. Then $\Omega_1 = I + \sum_{i=1}^{k} \lambda_i A_i A_i' = I + A\Gamma A'$ where $A = [A_1 \ldots A_k]$ and Γ is a diagonal matrix with diagonal elements equal to $\lambda_i I_{b_i}$, $i = 1, 2, \ldots, k$.

5. EXAMPLES

5.1 *Data Set 3*

We analyse the difference in conductance, maximum minus onset. Table 1 gives the differences. There are clearly difficulties as there were many zero differences and several values which appear to be 'outliers'; in particular note cases 3, 13 and perhaps 23.

Clearly any analysis of these data is problematic. However, we illustrate our approach using these data because it does highlight the possible insights and simplifications that can be achieved. The influence of outliers on the process is also made clear.

The data set has a simple treatment structure and we begin with a multivariate regression, using $\mathrm{E}(Y) = DT$ where D is a 24×2 matrix indicating the treatment group for each subject. The matrix of means, T is 2×6. Two preliminary analyses are given in Table 2, where the full data are used, and observations 3 and 13 are removed. The changes in estimates of the mean responses over time leave one a little uneasy, but observations 3 (time 1) and 13 are very different from the remainder. What is more striking is the change in the correlation structure and the variances. The variances are much more stable once the outliers have been removed (not surprisingly). The correlation structure, however, highlights the importance of examining the data not only for outliers but for second order structure, that is patterns in the correlations. It appears that at time 1 we have a larger variance than at the other times, which are reasonably uniform, that time 1 is essentially uncorrelated with the other time points and that the

Table 1

Differences in conductance

	Exposure					
Subject	1	2	3	4	5	6
1	1.8	1.4	1.3	1.3	0.8	0.6
2	0.0	0.0	0.0	0.0	0.0	0.0
3	14.1	0.0	2.4	0.0	0.0	0.0
4	0.0	0.7	0.7	0.9	0.0	1.6
5	0.2	0.1	0.2	0.0	0.0	0.0
6	1.8	0.0	1.6	0.0	0.0	0.0
7	0.0	0.0	0.0	0.0	0.0	0.0
8	0.3	0.0	0.0	0.0	0.0	0.0
9	2.6	0.5	0.0	0.0	0.0	0.2
10	0.9	0.3	0.0	0.0	0.0	0.0
11	0.7	0.0	0.0	0.0	0.0	0.0
12	0.5	0.3	1.2	0.9	0.3	0.0
13	8.9	6.0	0.0	0.0	3.6	0.0
14	0.0	0.0	0.0	0.0	0.0	0.0
15	1.0	0.0	0.0	0.0	0.6	0.0
16	0.1	0.0	0.0	0.0	0.0	0.0
17	1.8	1.1	0.2	0.0	0.0	0.0
18	0.6	0.7	0.0	0.0	0.6	0.4
19	0.0	0.0	0.0	0.0	0.0	0.0
20	1.0	3.2	1.8	1.9	2.0	3.3
21	3.0	1.9	0.0	0.2	0.0	0.0
22	1.5	3.3	1.9	2.5	1.8	2.0
23	5.3	0.0	0.0	0.0	0.5	0.0
24	0.7	0.5	0.0	0.3	0.0	0.0

remaining time points exhibit equi-correlation (perhaps not surprising in

Table 2

Multivariate regression analysis

(a) All the data

$\hat{T}$

1.9083	0.2750	0.6167	0.2583	0.0917	0.2000
1.9917	1.3917	0.3250	0.4083	0.7583	0.4750

Sample correlations and variances

11.2308					
0.3430	1.8833				
0.4021	0.3217	0.5927			
-0.1262	0.4751	0.6894	0.4772		
0.3579	0.9098	0.3433	0.4486	0.6772	
-0.1288	0.4401	0.5976	0.8451	0.4292	0.6738

(b) Observations 3,13 omitted

$\hat{T}$

0.8000	0.3000	0.4545	0.2818	0.1000	0.2182
1.3636	0.9727	0.3545	0.4455	0.5000	0.5182

Sample correlations and variances

1.6433					
0.1485	0.9091				
0.0880	0.7357	0.4727			
0.0048	0.8708	0.8560	0.5122		
0.1075	0.7885	0.7874	0.8680	0.3040	
-0.0710	0.8142	0.7275	0.8421	0.8330	0.7266

view of the number of zeros in the differenced data). This structure indicates that perhaps the initial exposure in the experiment was very different from the remainder, which were more or less uniform. The means exhibit similar trends. These remarks relate to the question of acclimatisation and given the structure it seems appropriate to test hypotheses in groups.

For each treatment group we test equality of the means of the last five differences, and if retained we test the hypothesis that the first mean is greater than the common mean of the last five differences. Because of the simple nature of the covariance structure, the initial tests of equality are F-tests, while the tests concerning the first mean and the remainder are Behrens-Fisher type tests, that is involve testing means with unequal variances. If one was to assume equal variances for all six measurement differences, t-tests would be appropriate. In fact given the robustness of the t-test to unequal variances this is not a bad approach.

Table 3

Tests of hypotheses for Data Set 3

Test of equality of last five mean differences		
Group	F-statistic	d.f.
1	0.32	4,18
2	1.10	4,18

Test first mean > common mean of last five exposures	
Group	Behrens-statistic
1	1.20
2	1.83

Test of equality of groups		
Time	t-statistic	d.f.
1	-1.03	20
2-6	-0.97	20

The other question of interest concerns the difference in the groups. These tests can be conducted using t-tests.

The results of these tests are given in Table 3. For both groups, the last five mean differences do not differ significantly. Testing that the mean difference at time 1 is greater than the common mean for the remaining times produces non-significant results for group 1 but for group 2 there is some evidence that there may be some acclimatisation. Testing for differences between groups is inconclusive.

5.2 *Data Set 4*

We found little evidence of correlation between CO_2 expiration and anxiety score during phobic talk or overbreathing.

The questions of interest to the researchers all relate to changes in CO_2 expiration and hence we analyse differences. The data then reflect the changes when subjects move from state to state, for example from rest to breathing; this is denoted by B-T in what follows, with similar notation for other movements. Because of the nature of the experiment, with treatments applied at various times, we expect both the mean differences and the variances and covariances to exhibit a good deal of structure. Our aim is to quantify that structure and provide a meaningful analysis of the data.

We use repeated measures and differences interchangeably in what follows. In the notation of Section 2, we have $n = 41$ experimental units, $p = 10$ repeated measures, $r = 2$ treatments, and $m = 21$ missing values, made up of $(0, 0, 0, 0, 0, 1, 3, 8, 8, 1)$ for the 10 repeated measures. Note that differencing increases the effective number of missing values.

Our model is given by (4) with F a 410×21 matrix,

$$D = \begin{bmatrix} 1_{22} & 0_{22} \\ 0_{19} & 1_{19} \end{bmatrix} \qquad X = I_{10} \otimes D$$

and $\Omega_1 = I_n$, so that units are assumed independent.

We begin by assuming the differences on each individual are uncorrelated and have equal variance; that is we set $\Omega_2 = I_p$. We fitted the linear model (4) and found a residual log-likelihood of -346.72. We used this fit to produce Table 4 which gives the estimated variances (down the diagonal) and correlations (in the lower triangular part), adjusted to take into account the missing data. Various attempts were made to model the structure. For example, independence with unequal variances and the split-plot form which is essentially the analysis of variance approach were two that we tried. The second is certainly questionable given Table 4 and using the analysis of variance may lead to misleading results.

The margins of Table 4 are labelled by the differences in states and by + or 0. The latter two labels indicate whether or not the differenced data involves a change in state. For example, T-R is the difference between two different states, namely talk and rest, and so whenever T-R appears in a margin, it is also associated with +. Differences which involve the same state, for example T-T talk minus talk, have 0 in the margin. Both

Table 4

Variances/correlations for Data Set 4

T-R	+	2.68									
T-T	0	0.01	5.15								
T-T	0	-0.02	0.15	3.18							
R-T	+	-0.66	-0.31	-0.41	2.64						
R-R	0	-0.34	-0.32	-0.49	0.44	3.60					
B-R	+	0.09	-0.13	-0.03	-0.23	-0.18	21.04				
B-B	0	0.05	0.06	-0.19	-0.12	0.13	-0.02	1.93			
B-B	0	-0.06	-0.35	-0.42	0.34	0.28	0.24	0.12	0.88		
R-B	+	-0.21	0.32	0.06	-0.06	0.01	-0.28	0.28	-0.06	6.94	
R-R	0	-0.05	0.11	0.28	0.07	-0.07	-0.50	-0.48	-0.48	-0.04	17.06
Change		+	0	0	+	0	+	0	0	+	0
Diff.		T-R	T-T	T-T	R-T	R-R	B-R	B-B	B-B	R-B	R-R

margins, rows and columns, are labelled so that the correlation structure adopted can be discussed below.

We progress in stages. Firstly we examine the estimated variances given in Table 4. The estimated variances appear stable until the B-R, breathing minus rest, period, where large variation is evident. The variation under breathing is stable until *there is a change* to rest where it increases again. We model the variances to reflect the fact that changes from one state to another do result in changes in variation. We model the variances to take these changes into account using

$$\text{variances} = \sigma^2(1, 1, 1, 1, 1, \lambda_1, \lambda_2, \lambda_2, 1, \lambda_1) \tag{6}$$

where the last two parameters conform with previous ones because of obvious reduction in model complexity. In the implementation of the estimation procedure, $\lambda_i = \exp\theta_i$ to ensure non-negative estimates of variances. We fitted this structure for the variances, the differences again assumed uncorrelated, and the residual log-likelihood increased to -277.54. This is a substantial increase for the addition of two parameters.

The correlation structure appears more complex. Interesting subsets of correlations correspond to the pairs $(+,+)$, $(0,0)$ and $(+,0)$. These correspond to correlations between two differences both of which are changes in state, both of which have no change in state and, one of each, respectively. While the patterns are not entirely consistent, the $(+,+)$ group has a predominance of negative correlations with an average of -0.225, the $(0,0)$ group also has some large negative correlations, with an average correlation of -0.112, while the last group, $(+,0)$ has a predominance of small correlations with an average of -0.018. The model we propose for the correlation structure is

$$R = I + \theta_3 R_{++} + \theta_4 R_{00} \tag{7}$$

where the matrices R_{++} and R_{00} are indicator matrices with unit elements for the $(+,+)$ and $(0,0)$ positions respectively. Our full covariance structure is

$$\sigma^2 \Omega_2 = \sigma^2 LRL$$

where L is a diagonal matrix of standard deviations, the square root of the variances modelled in (6).

Fitting this correlation structure together with the variance model resulted in a residual log-likelihood of -267.5, a reasonable increase over the previous model.

Table 5

Residual log-likelihoods for covariance structure

Model	log-likelihood	χ^2	df
$\Omega_2 = I_{10}$	-346.72		
$\Omega_2 = L^2$	-277.54	138.36	2
$\Omega_2 = LRL$	-267.50	20.08	2
$\Omega_2 = LR^*L$ $(\theta_3 = \theta_4)$	-269.55	4.10	1

Table 5 summarises the REML log-likelihoods, approximate chi-square statistics and degrees of freedom for testing the three models discussed

Table 6

Difference between Panic and Control Groups

Panic Group		Control Group	t-statistic	
T-R	-0.58	-0.91	-0.52	T-R
T-T	-0.74	-1.41	-1.08	T-T
T-T	-0.58	-0.18	0.64	T-T
R-T	0.68	0.68	0.00	R-T
R-R	0.53	0.50	0.04	R-R
B-R	-15.68	-11.98	2.68	B-R
B-B	-2.82	-2.58	0.45	B-B
B-B	-1.63	-1.03	0.78	B-B
R-B	8.19	5.04	-4.17	R-B
R-R	7.21	5.57	-1.19	R-R

above in turn and testing $\theta_3 = \theta_4$. Using the approximate chi-square tests, we arrive at the model based on (6) and (7).

Table 6 contains the estimated effects for each difference for each of the two treatment groups. An approximate t-statistic is provided for the test of no difference between the panic and control groups. We see that the only important changes occur when the subjects move from rest to breathing, and here the panic group exhibits a greater reduction in CO_2 expiration, and from breathing to rest where the panic group has the larger increase in CO_2 expiration. These results mirror the changes in variance and we conclude that although variation increases at these event times, deep breathing helps the panic group in terms of reduction of CO_2, but that a more rapid rise occurs when these subjects move from breathing to rest.

5.3 *Apple trees*

This experiment was carried out to obtain yield production curves for three varieties of apples grown under 4 training systems (pruning schemes). In particular, the question of which training scheme provided the best early yield was of interest. The factors in the experiment together with their levels are given in Table 7. The experiment was carried out in the field

exactly as shown in Table 8; there is little evidence of statistical advise with regard to randomization and replication. The additional complication is that there were three plantings, from 1969 to 1971. Thus we consider tree age rather than years as the time variable, but we allow for year effects in the modelling, as described below.

Table 7

Treatment structure for apple experiment

Rootstock	Variety	Training System
MM106 (106)	Jonathan (J)	Vase (V)
Northern spy (NS)	Delicious (D)	Hawkes Bay (H)
MM102 (102)	Granny Smith (G)	Palmette (P)
Seedling (S)		Central Leader (C)

The 'design' is a split plot with rows representing the whole plots, with whole plot treatments being the rootstock by variety combinations. The plots within each row were randomly allocated a training system. Plot yields were measured from 1972 to 1987. The variety delicious was incompatible with rootstock MM102 and all these trees died soon after planting. There was one other plot in which the trees died after 2 years and we have discarded these data.

Many fruit trees are biennial bearing, that is one year on, one year off. We analyse two yearly totals. For trees planted in 1969 we have $p = 8$ two-yearly yields, for trees planted in 1970 we have $p = 7$ two-yearly yields plus a single years yield and for the plots planted in 1971 we have $p = 7$ two-yearly yields.

There were large trends across the field (4.11 hectares). Because of this and the restricted randomisation, we imposed an additional blocking factor as indicated in Table 8. These blocks align essentially with the rootstock by replicate strata, with the exception of blocks in rows 13 to 15. The division into two blocks was justified on the basis that the two sections of the same rows were planted in different years. It also simplifies the analysis.

The full mean model, rootstock by variety by training scheme together with up to 8 repeated measures was examined at each time point. Based on

Table 8

Field layout for apple experiment

Row	Rootstock	Variety	Plot 1 2 3 4 5 6 7 8	
2	106	G	V P C H C P V H	
3	106	D	C H V P V P H H	
4	106	J	C P H V C H P V	
5	S	G	H V C P	Planted 1969
6	S	D	H V P C	
7	S	J	P H V C	
9	102	J	V P C H V H C P	
10	102	D	V H P C P H V C	
11	102	G	C P V H H V C P	
13	NS	J	P V H C P C V H	
14	NS	G	H C V P H V P C	
15	NS	D	V H P C C P H V	
16	106	G	P H C V	
17	106	D	H C V P	
18	106	J	P C H V	Planted 1970
19	S	D	V P C H	
20	S	G	C V P H	
21	S	J	H C V P	
22	NS	G	P H C V	Planted 1971
23	NS	J	H V C P	
24	NS	D	C V P H	
25	S	J	P C H V	
26	S	D	V C H P	
27	S	G	C P H V	Planted 1970
29	102	D	C V H P	
30	102	J	V H P C	
31	102	G	P C H V	

these preliminary analyses and information provided by the researcher, we considered the reduced model of rootstock plus variety by training system.

The analysis proceeds by considering yield at specific tree ages, treating the three plantings as cohorts, and including year effects, to allow for the possibility of differences due to the year of planting and subsequent conditions. This is consistent with the aims of the experiment, namely to produce yield curves as a function of tree age. The model for each of the three sets ($j = 1, 2, 3$) in matrix form is

$$Y_j = D_j T + W_j \phi_j' + A_j B_j + E_j$$

where D_j, W_j and A_j are design matrices for treatments, years and blocks respectively and T, ϕ and B_j are the corresponding parameters. In vector form we have

$$y_j = X_j \tau + F_j \phi_j + Z_j \beta_j + e_j$$

for $j = 1, 2, 3$ where τ is the vector of common treatment effects, ϕ_j is the vector of year effects for set j, and β_j is the vector of block effects for set j. The design matrices X_j, F_j and Z_j are formed using the direct product as in Sections 2 and 4. The mean vector and covariance matrix are

$$\mathrm{E}(y_j) = X_j \tau + F_j \phi_j \qquad \mathrm{var}(y_j) = \sigma^2 \Omega_{2j} \otimes \Omega_{1j}$$

where $\Omega_{1j} = I_{n_j} + \lambda A_j A_j'$, reflecting the blocking structure, $\Omega_{2j} = K_j \Omega_2 K_j'$, the K_j being incidence matrices which extract the appropriate portion of Ω_2.

If $y = (y_1', y_2', y_3')'$, then

$$y = X\tau + F\phi + Z\beta + e$$

where $F = \mathrm{diag}(F_j)$, $Z = \mathrm{diag}(Z_j)$ and

$$X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}, \quad \phi = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}$$

Furthermore

$$\mathrm{E}(y) = X\tau + F\phi, \qquad \mathrm{var}(y) = \sigma^2 \mathrm{diag}(\Omega_{2j} \otimes \Omega_{1j}) \tag{8}$$

Having fixed the mean model we turn to the determination of the temporal covariance structure. We initially assume $\Omega_2 = I$. Fitting (8) with the block structure in Ω_{2j} we obtained the SSP/correlation matrix given in Table 9. The temporal error process is non-stationary and transforming the data by logarithms is too severe. Rather than transform we model the non-stationarity using a model proposed by GEARY [6]. The model allows for increasing variances and covariances. The GEARY model is

$$\Omega_2 = \rho J + V_G(\xi, \zeta)$$

where J is a matrix of ones, and V_G is for example when $p = 3$

$$V_G = \frac{1}{1-\xi\zeta^2}\begin{bmatrix} 1 & & \\ \zeta & 1-\xi\zeta^2+\zeta^2 & \\ \zeta^2 & \zeta(1-\xi\zeta^2+\zeta^2) & 1-\xi\zeta^2+\zeta^2-\xi\zeta^4+\zeta^4 \end{bmatrix}$$

The parameter ρ was very small in our application and so we fitted a model without the term involving ρ. REML estimates of the remaining parameters were $\hat{\zeta} = 0.548$, $\hat{\xi} = -212$ and the block variance to error variance ratio, $\hat{\lambda} = 0.47$.

Table 9

Variances/correlations for Apple Data Set

90							
0.26	2368						
0.07	0.61	4551					
0.13	0.63	0.69	9353				
0.16	0.31	0.31	0.65	9036			
-0.29	0.10	0.03	0.07	0.29	12640		
0.04	0.23	0.13	0.47	0.64	0.26	13526	
0.08	0.39	0.20	0.61	0.60	0.23	0.75	9491

Taking this covariance structure we proceed with modelling to establish which training scheme produces the best early yield. The variety by

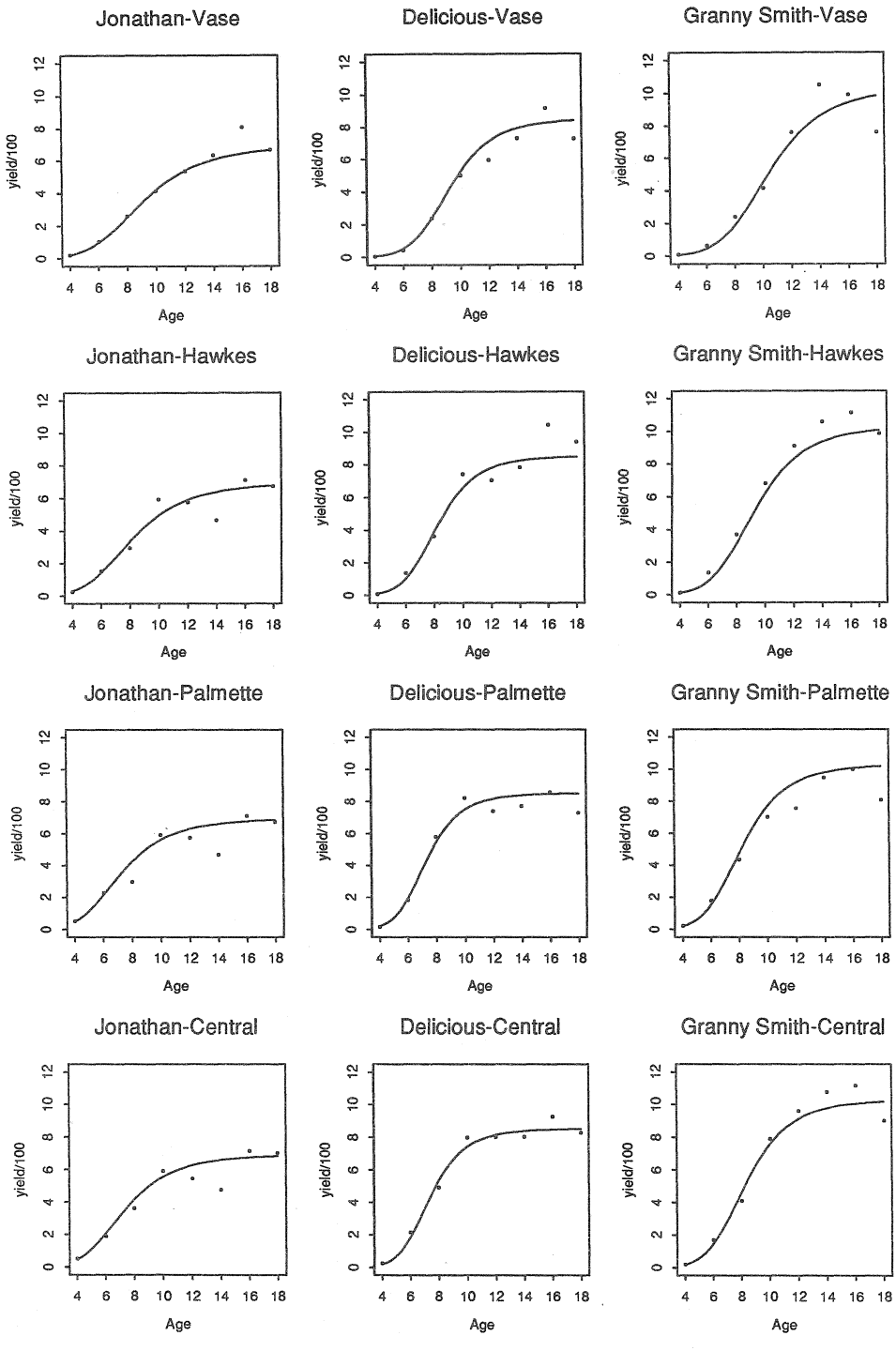

Figure 1. Variety by training scheme effects and fitted yield curves

Table 10

REML estimates of log-logistic curve

Variety	Training scheme	D_{ij}	α_{ij}	m_{ij}	F statistic
Jonathan	Vase	818	4.01	2.30	2.2
	Hawkes	675	4.80	2.07	7.0
	Palmette	717	4.33	1.97	4.1
	Central Leader	695	3.88	2.03	6.1
Delicious	Vase	841	5.56	2.26	3.2
	Hawkes	936	6.09	2.13	6.9
	Palmette	811	7.14	1.95	2.3
	Central Leader	885	5.58	2.02	2.2
Granny Smith	Vase	987	6.18	2.33	8.0
	Hawkes	1143	5.25	2.22	1.3
	Palmette	933	5.13	2.11	3.4
	Central Leader	1079	5.40	2.15	3.2

training scheme effects over time are given in Figure 1 together with the fitted profiles we now discuss. We consider the mean yield curve

$$\eta_{ijt} = \frac{D_{ij}}{1 + e^{-\alpha_{ij}(\log t - m_{ij})}}$$

where i indexes the variety and j indexes the training scheme. The parameters D_{ij} are the asymptotes, α_{ij} are shape parameters and m_{ij} are log half-effect parameters. Table 10 gives the parameter estimates and approximate individual F statistics for testing the fit. There are some large F values indicating lack of fit. However some patterns are clear.

- The variety factor affects D, α and m,
- the training scheme affects only m, and
- there is no obvious interaction of variety and training scheme for the m parameters so that we have the additive model $m_{ij} = m_i + m_j$, additive due to variety and to training scheme.

Our reduced model is

$$\eta_{ijt} = \frac{D_i}{1 + e^{-\alpha_i(\log t - m_i - m_j)}}$$

Table 11

REML estimates of reduced log-logistic curve

	Jonathan	Delicious	Granny Smith
D_i	697 (34)	853 (32)	1033 (34)
α_i	4.39 (0.24)	6.27 (0.34)	5.53 (0.25)
$m_{ij} = m_i + m_j$			
	Jonathan	Delicious	Granny Smith
Vase	2.21 (0.04)	2.23 (0.03)	2.35 (0.02)
Hawkes Bay	2.10 (0.04)	2.11 (0.03)	2.23 (0.02)
Palmette	1.98 (0.04)	1.99 (0.02)	2.11 (0.02)
Central Leader	1.99 (0.04)	2.00 (0.02)	2.12 (0.02)

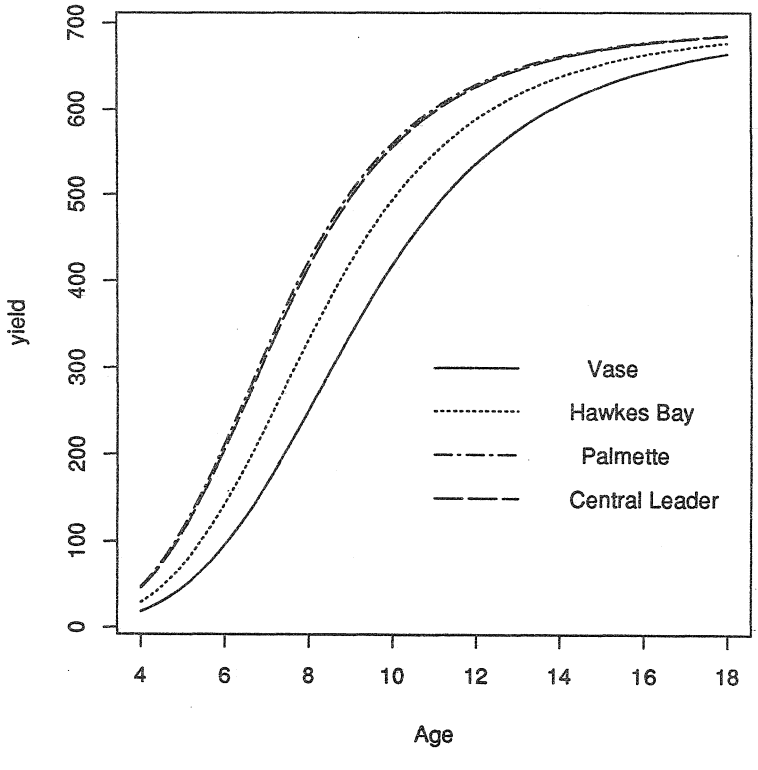

Figure 2. Estimated yield curves for Jonathan apples

and an overall approximate F statistic is 4.50. This indicates that there is significant variation still present. This type of problem arises in many contexts and is the subject of current research. The estimates for the reduced model are given in Table 11 and plots of the fitted model are given in Figure 1. The unaccounted variation apparently occurs towards the end of the study. Nonetheless, the model does provide important information that can be used by the researcher. Figure 2 provides the typical information deduced, that palmette and central leader give the greatest initial yield with the traditional vase being the slowest in yield uptake. This applies to all variety types.

6. DISCUSSION

Specific discussion relating to the examples occurred both after our talk and during the talks on analysis of the supplied data sets.

Data set 3 has many zero elements after differencing, and the normality assumption was questioned.

The choice of the covariance structure for Data set 4 is to some extent controversial, but the aim of data analysis is to determine as much structure as possible and we feel that the structure chosen reflects where important changes are occurring. A question was asked regarding the calculation of the sample covariance matrix in Table 4 in view of the missing data. An approximate allowance was made for incompleteness in that the divisors were the actual degrees of freedom at each 'time'.

For the apple data set there was some reservation (by the speaker!) that the covariance structure was entirely appropriate. The biennial bearing problem and two year totals was queried privately, and as a result the full data are to be re-analysed. The main discussion point centred on the modelling of the variety by training scheme profiles. The model chosen did not fit the data well and the question was asked whether a parametric model was necessary. Given the aim of the experiment was to determine yield curves and in particular to examine how the training scheme affected the rate at which yield increased in the early years of development, a parametric approach such as the one adopted was necessary. Furthermore, while the model is not entirely satisfactory, it does provide a simple summary of the main features of the data, for example as given in Figure 2.

REFERENCES

[1] COOPER, D. M. and THOMPSON, R. (1977). A note on the estimation of the parameters of the autoregressive-moving average process. *Biometrika* **64**, 625-628.

[2] CULLIS, B. R. and GLEESON, A. C. (1991). Spatial analysis of field experiments - an extension to two dimensions. *Biometrics*, to appear.

[3] CULLIS, B. R. and MCGILCHRIST, C. A. (1990). A model for the analysis of growth data from designed experiments. *Biometrics* **46**, 131-142.

[4] CULLIS, B. R. and VERBYLA, A. P. (1991). Nonlinear regression and time dependent covariates in repeated measures experiments. *Australian Journal of Statistics*, to appear.

[5] DIGGLE, P. J. (1988). An approach to the analysis of repeated measurements. *Biometrics* **44**, 959-971.

[6] GEARY, D. N. (1989). Modelling the covariance structure of repeated measurements. *Biometrics* **45**, 1183-1196.

[7] HARVILLE, D. A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association* **72**, 320-340.

[8] HOUTMAN, A. and SPEED, T. P. (1984). The analysis of multistratum designed experiments with incomplete data. *Australian Journal of Statistics* **26**, 227-246.

[9] LAIRD, N. M. and WARE, J. H. (1982). Random effects models for longitudinal data. *Biometrics* **38**, 963-974.

[10] MARTIN, R. J. (1990). The use of time series models and methods in the analysis of agricultural field trials. *Communications in Statistics - Theory and Methods* **19**, 55-81.

[11] PATTERSON, H. D. and THOMPSON, R. (1971). Recovery of interblock information when block sizes are unequal. *Biometrika*, **58**, 545-554.

[12] VERBYLA, A. P. (1988). Analysis of repeated measures designs with changing covariates. *Biometrika* **75**, 170-172.

[13] VERBYLA, A. P. (1990). A conditional derivation of residual maximum likelihood. *Australian Journal of Statistics* **32**, 227-230.

[14] VERBYLA, A. P. (1992). Modelling variance heterogeneity: residual maximum likelihood and diagnostics. *Submitted for publication.*

[15] VERBYLA, A. P. and CULLIS, B. R. (1990). Modelling in repeated measures experiments. *Applied Statistics* **39**, 341-356.

[16] VERBYLA, A. P. and CULLIS, B. R. (1992). The analysis of multistratum and spatially correlated repeated measures data. *Biometrics*, to appear.

Department of Statistics,
University of Adelaide,
GPO Box 498,
Adelaide, S. A. 5001
and
Agricultural Research Institute,
Wagga Wagga, N. S. W., 2650