RT-MAE 9841

ANALYSING LONGITUDINAL DATA
VIA NONLINEAR MODELS IN
RANDOMIZED BLOCK DESIGNS

by

Paulo J. Ogliari
and
Dalton F. Andrade

Palavras-Chave: Repeated measures; Growth curve; Experimental designs; Mixed-effects models.

Classificação AMS: 62J02, 62K10, 62P99
(AMS Classification)

- Novembro de 1998 -

Analysing Longitudinal Data Via Nonlinear Models in Randomized Block Designs

Paulo J. Ogliari and Dalton F. Andrade[1]

Abstract

Two ways of introducing block effects into nonlinear models are considered. One enters the block effects in a linear fashion and the other in a nonlinear fashion. In both cases block effects can be considered as fixed or random. Additionally to the block effects, we discuss the incorporation of different covariance structures into the model to take into account the longitudinal nature of the data. Existing methods for fitting nonlinear models for data from completely randomized designs are shown to be appropriate for this setup. A numerical growth curve example for eucalyptus trees is presented.

Key Words: Repeated measures; Growth curve; Experimental designs; Mixed-effects models.

1. Introduction

Longitudinal data are very frequently in many research areas, in particular in agricultural and biological studies. They occur when two or more observations of a response variable are obtained at different instants for each subject under investigation. One example would be the study conducted at Klabin Fabricadora de Papel e Celulose SA do Paraná, Brazil, where the solid volume with bark was evaluated for each of 16 subjects (groups of eucalyptus trees)

[1]Paulo J. Ogliari is Associate Professor, Department of Computer and Statistics, University of Santa Catarina, P.O. 476, 88040-900 Florianópolis - Santa Catarina - Brazil (E-mail:ogliari@inf.ufsc.br). Dalton F. Andrade is Assistant Professor, Department of Statistics, University of São Paulo, P.O. 20570, 05508-900 São Paulo - Brazil, (E-mail:dfa@ime.usp.br)

at ages 3, 4, 5 and 9 years. Four subjects were randomly assigned to one of four treatments (2 species $\times$ 2 spacings) according to a complete randomized block design. The data are presented in Table 5 (Appendix A). The reader is referred to Goldstein (1979), Crowder and Hand (1990) or Lindsey (1993) among others, for a broadly caracterization of these type of data. In this paper our main goal is to model longitudinal data obtained from randomized block designs when the behavior of the response variable along one ordered dimension (as time, for instance) is well represented by nonlinear functions. Particularly, studies to describe agricultural growth processess call for a mean response function which is inherently nonlinear in the parameters.

Let $\mathbf{y}_{ji} = (y_{ji1}, y_{ji2}, ..., y_{jip_{ji}})^t$ be the $(p_{ji} \times 1)$ vector of the responses for subject i in block j $(j = 1, 2, ..., b$ and $i = 1, 2, ..., n_j)$, $\mathbf{f}$ be a nonlinear vector valued function, $\mathbf{X}_{ji}$ be a $(p_{ji} \times r)$ known matrix of explanatory (or independent) variables and $\boldsymbol{\alpha}$ be an $(m \times 1)$ unknown vector of location parameters. Ignoring block effect, the simplest nonlinear model to represent the relationship between $\mathbf{y}_{ji}$ and the explanatory variables would be

$$\mathbf{y}_{ji} = \mathbf{f}(\mathbf{X}_{ji}, \boldsymbol{\alpha}) + \boldsymbol{\epsilon}_{ji} \tag{1.1}$$

with $\boldsymbol{\epsilon}_{ji}$, $(p_{ji} \times 1)$, independent random error vectors whose elements are also independents with mean zero and variance σ^2. Under this setup, the covariance structure induced for the response vector $\mathbf{y}_{ji}$ is $\sigma^2 \mathbf{I}_{p_{ji}}$, where $\mathbf{I}_{p_{ji}}$ is the identity matrix of order p_{ji}. For our data set, a suitable vector function $\mathbf{f} = (f_1,\ f_2,\ f_3,\ f_4)^t$ for the $ji-th$ subject assigned to treatment k $(k = 1, 2, 3, 4)$ would be the Gompertz growth curve model

$$f_s(\mathbf{X}_{ji}, \boldsymbol{\alpha}) = \alpha_{0k} exp\{-exp[-\alpha_{1k}(x_{jis} - \alpha_{2k}/\alpha_{1k})]\} \tag{1.2}$$

$s = 1, 2, 3, 4$ corresponding to the ages, where $\mathbf{X}_{ji} = (x_{ji1}\ x_{ji2}\ x_{ji3}\ x_{ji4})^t = (3\ 4\ 5\ 9)^t$, $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1^t\ \boldsymbol{\alpha}_2^t\ \boldsymbol{\alpha}_3^t\ \boldsymbol{\alpha}_4^t)^t$ with $\boldsymbol{\alpha}_k = (\alpha_{0k}\ \alpha_{1k}\ \alpha_{2k})^t$, α_{0k} representing an asymptote (maximum volume), α_{1k} a constant of proportionality, α_{2k}/α_{1k} corresponding to the age at which the growth rate reaches its maximum. This model was used with success in the fit of volume

Table 1: Sample variances (main diagonal), covariances (above the main diagonal) and correlations (below the main diagonal).

Age (years)	Age (years) 3	4	5	9
3	286.1092	404.5760	574.7248	910.4979
4	0.8547	783.0833	1012.7649	1673.4248
5	0.9110	0.9704	1390.8271	2274.7045
9	0.7969	0.8853	0.9030	4562.2151

with age for *Pseudotsuga menziesii* (Mrb) Franco, Nokoe (1980). Other examples for the f function would be the Exponential, Richard, Von Bertalanffy and Logistic models described in Ratkowsky (1983), for instance.

Two drawbacks of this formulation must be pointed out. One is that it does not incorporate possible dependence among the observations taken on the same subject. The second one is that in longitudinal studies it is quite commom to have the variances varying along the ordered dimension. Singer and Andrade (1994) present one example in the field of Animal Nutrition where the variances decrease with time. Table 1 shows the sample covariance-correlation matrix, corrected for treatment effect, obtained for our data. One can see that the variances increase with age and the correlations between observations taken at any two instants are high and, in general, decrease as the distance between them increases.

Addicionally, when data are collected from randomized block designs and block effects are considered random, as in most of the practical situations, dependence among the observations taken on the subjects belonging to the same block need also to be considered in the model. Therefore, it is important to have models that allow for both dependences (within and between subjects) and also for heterocedasticity in their formulations.

In section 2 we introduce the block effects in the model and in section 3 we present several variance-covariance structures. Statistical inference for these models is discussed in section 4 and in section 5 we analyse the data presented in Table 5. In the last section we present some additional comments and conclusions.

2. Modelling the Block Effects

In general, blocking is only used to control sample units' heterogeneity and it is not expected significant block x time and block x treatment interactions. Therefore, paralleling the work by Andrade and Singer (1998), we introduce the block effects in the model as an extra factor in a linear fashion

$$\mathbf{y}_j = \mathbf{f}(\mathbf{X}_j, \boldsymbol{\alpha}_j) + \mathbf{1}_{p_j}\beta_j + \boldsymbol{\epsilon}_j \quad j = 1, 2, ..., b, \tag{2.1}$$

where $\mathbf{y}_j = (\mathbf{y}_{j1}^t\ \mathbf{y}_{j2}^t\ ...\ \mathbf{y}_{jn_j}^t)^t$, $\mathbf{X}_j = (\mathbf{X}_{j1}^t\ \mathbf{X}_{j2}^t\ ...\ \mathbf{X}_{jn_j}^t)^t$, $\boldsymbol{\alpha}_j = \mathbf{A}_j\boldsymbol{\alpha}$ with $\mathbf{A}_j$ a $(q_j \times m)$ known matrix, $\boldsymbol{\epsilon}_j = (\boldsymbol{\epsilon}_{j1}^t\ \boldsymbol{\epsilon}_{j2}^t\ ...\ \boldsymbol{\epsilon}_{jn_j}^t)^t$ with $\mathbf{y}_{ji}^t$, $\mathbf{X}_{ji}$, $\boldsymbol{\alpha}$ and $\boldsymbol{\epsilon}_{ji}$ as in (1.1), $\mathbf{1}_{p_j} = (1\ 1\ ...\ 1)^t$ of dimension $(p_j \times 1)$, $p_j = p_{j1} + p_{j2} + ... + p_{jn_j}$ and β_j is the effect of block j such that:

1. $\beta_1 + \beta_2 + ... + \beta_b = 0$, for fixed blocks and
2. $\beta_1, \beta_2, ..., \beta_b$ *i.i.d.* random variables with mean zero and variance σ_b^2, and independent of the error terms, for random blocks.

In the first case, the block effects do not enter in the covariance structure of $\mathbf{y}_j$ but, on the other hand, in the second case where the blocks are random, as it is in most of the practical applications (see Neter et al (1990), for instance), the variance component σ_b^2 enter in the covariance matrix $\boldsymbol{\Sigma}_j$ of $\mathbf{y}_j$ in such a way that

$$\boldsymbol{\Sigma}_j = \sigma_b^2\mathbf{J}_{p_j} + \mathbf{I}_{n_j} \otimes (\sigma^2\mathbf{I}_{p_{ji}}) \quad j = 1, 2, ..., b \text{ and } i = 1, 2, ..., n_j \tag{2.2}$$

where $\mathbf{J}_{p_j}$ is a $(p_j \times p_j)$ matrix of one's. This formulation induces an uniform covariance structure for the observations taken on the same block, which, as pointed out in section 1, it is not expected for these type of data. In particular, it is also not expected that the correlation between two observations taken on the same subject be equal to the correlation between two observations taken on different subjects in the same block. Therefore, it is necessary to introduce a more general specification for $\boldsymbol{\Sigma}_j$.

Let $\boldsymbol{\Sigma}_{ji}$ be the $(p_{ji} \times p_{ji})$ covariance matrix of $\mathbf{y}_{ji}$ as if the data were collected from a completely randomized design, then

$$\boldsymbol{\Sigma}_j = \sigma_b^2 \mathbf{J}_{p_j} + \mathbf{I}_{n_j} \otimes \boldsymbol{\Sigma}_{ji} \quad j = 1, 2, ..., b \text{ and } i = 1, 2, ..., n_j \tag{2.3}$$

and different specifications for $\boldsymbol{\Sigma}_j$ can be obtained from different structures for $\boldsymbol{\Sigma}_{ji}$.

Block effects could also enter the model in a nonlinear fashion such as

$$\mathbf{y}_j = \mathbf{f}(\mathbf{X}_j, \boldsymbol{\alpha}_j^*) + \boldsymbol{\epsilon}_j \quad j = 1, 2, ..., b, \tag{2.4}$$

with $\boldsymbol{\alpha}_j^* = \boldsymbol{\alpha}_j + \mathbf{b}_j\beta_j$, $\mathbf{b}$ a known $(q_j \times 1)$ vector, and $\boldsymbol{\alpha}_j$ and β_j as above. A natural value for $\mathbf{b}$ would be $\mathbf{1}_m = (1\ 1\ ...\ 1)^t$ which associates the block effects to all the location parameters. For random blocks, paralleling the nonlinear mixed-effects models setup (see Vonesh and Carter (1992), Vonesh (1992) and Vonesh and Chinchilli (1997), for instance), it is convenient to consider a first-order Taylor's serie expansion of the above function expanded about 0, the average of the random effects β_j,

$$\mathbf{y}_j = \mathbf{f}(\mathbf{X}_j, \boldsymbol{\alpha}_j) + \mathbf{Z}_j(\boldsymbol{\alpha}_j)\beta_j + \boldsymbol{\epsilon}_j \quad j = 1, 2, ..., b, \tag{2.5}$$

where $\mathbf{Z}_j(\boldsymbol{\alpha}_j) = \left[\partial \mathbf{f}(\mathbf{X}_j, \boldsymbol{\alpha}_j^*)/\partial \boldsymbol{\alpha}_j^{*t}|_{\boldsymbol{\alpha}_j^* = \boldsymbol{\alpha}_j}\right] \mathbf{b}$. This formulation induces the following covariance structure for $\mathbf{y}_j$

$$\boldsymbol{\Sigma}_j = \sigma_b^2 \mathbf{Z}_j(\boldsymbol{\alpha}_j)\mathbf{Z}_j^t(\boldsymbol{\alpha}_j) + \mathbf{I}_{n_j} \otimes \boldsymbol{\Sigma}_{ji} \quad j = 1, 2, ..., b \text{ and } i = 1, 2, ..., n_j. \tag{2.6}$$

Again, a dependence structure among all the observations belonging to the same block is modelled but, differently from (2.3), this structure depends also on the function $\mathbf{f}$ and on the location parameters $\boldsymbol{\alpha}$.

3. Modelling the Covariance Matrix Σ_{ji}

In this section we focus our discussion on modelling $\boldsymbol{\Sigma}_{ji}$, the covariance structure due to the repeated measurement. There are many models proposed in the literature. The two more extreme in terms of the number of parameters would be

$$\boldsymbol{\Sigma}_{\text{ji}} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1p_{ji}} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2p_{ji}} \\ . & . & \dots & . \\ . & . & \dots & . \\ \sigma_{p_{ji}1} & \sigma_{p_{ji}2} & \dots & \sigma_{p_{ji}}^2 \end{bmatrix} \tag{3.1}$$

the so called unstructured model, where σ_s^2 is the variance at instant s ($s = 1, 2, ..., p_{ji}$) and σ_{su} is the covariance between the responses taken at instants s and u ($s \neq u$), and

$$\boldsymbol{\Sigma}_{\text{ji}} = \sigma^2 \begin{bmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ . & . & \dots & . \\ . & . & \dots & . \\ \rho & \rho & \dots & 1 \end{bmatrix} \tag{3.2}$$

the so called compound-symmetry structure, where σ^2 is the variance at any instant and ρ is the correlation between two observations taken at any two instants on the same subject. While structure (3.1) requires too many parameters ($= p_{ji}(1 + p_{ji})/2$), as the number of observations taken on the same subject increases, and consequently large sample size, structure (3.2) requires only two parameters, no matter the number of observations taken on the same subject. By the other hand, structure (3.2) assumes not only constant variances but also constant covariances; it does not happen with structure (3.1). As in general, we do not have large sample size and it is not expected structure (3.2) be appropriate for most longitudinal data, it is important to consider parsimonious structures like, for instance the heterogeneous AR(1) structure

$$\boldsymbol{\Sigma}_{\text{ji}} = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \dots & \sigma_1\sigma_p\rho^{p-1} \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \dots & \sigma_2\sigma_p\rho^{p-2} \\ . & . & \dots & . \\ . & . & \dots & . \\ \sigma_p\sigma_1\rho^{p-1} & \sigma_p\sigma_2\rho^{p-2} & \dots & \sigma_p^2 \end{bmatrix} \tag{3.3}$$

which requires $(p+1)$ parameters. Notice that, in the above structure all the subject are supposed to be observed at the same equally spaced instants. Wolfinger (1996), presents various alternatives structures which allow for different correlation patterns with a small number of parameters. Andrade and Helms (1986), discuss a linear structure where the covariance matrix is written as a linear combination of known matrices. Maybe, the simplest model for Σ_{ji} would be $\sigma^2\mathbf{R}_{ji}$, with $\mathbf{R}_{ji}$ a known positive definite matrix of order p_{ji}. For instance, in our example one possible value for $\mathbf{R}_{ji}$ would be

$$\mathbf{R}_{\mathrm{ji}} = \begin{bmatrix} 3 & 1 & 1 & 1 \\ 1 & 4 & 1 & 1 \\ 1 & 1 & 5 & 1 \\ 1 & 1 & 1 & 9 \end{bmatrix}. \tag{3.4}$$

This structure imposes increasing variances with time and decreasing correlations between observations taken at two instants as the distance between them increases, the pattern observed in our example.

4. Statistical Inference

For longitudinal data obtained from completely randomized designs, there are many methods in the literature to fit nonlinear models. In particular, under the random effects approach. These methods are mostly based on least square, maximum likelihood and bayesian estimation procedures. The reader is referred to Gallant (1987), Seber and Wild (1989), Gennings, Chinchilli and Carter (1989), Davidian and Giltinan (1995) and Vonesh and Carter (1997), among others, for details. In our case, where we have data from randomized block designs, the models proposed can be fitted throughout these methods modified to incorporate the block effects. We will focus our attention to the maximum likelihood estimation methods with data following a gaussian distribution.

Fixed blocks

When blocks are fixed, the subject-response vectors $\mathbf{y}_{ji}$ are normally independent distributed and the log-likelihood function can be written as

$$\begin{aligned} L(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}) &= (1/2)log(2\pi) \sum_{j=1}^{b} \sum_{i=1}^{n_b} p_{ji} - (1/2) \sum_{j=1}^{b} \sum_{i=1}^{n_b} log|\boldsymbol{\Sigma}_{ji}| \\ & -(1/2) \sum_{j=1}^{b} \sum_{i=1}^{n_b} [\mathbf{y}_{ji} - \mathbf{f}(\mathbf{X}_{ji}, \boldsymbol{\alpha}_j, \beta_j)]^t (\boldsymbol{\Sigma}_{ji})^{-1} [\mathbf{y}_{ji} - \mathbf{f}(\mathbf{X}_{ji}, \boldsymbol{\alpha}_j, \beta_j)] \quad (4.1) \end{aligned}$$

with $\mathbf{f}(\mathbf{X}_{ji}, \boldsymbol{\alpha}_j, \beta_j)$ given by the expected value of $\mathbf{y}_{ji}$ obtained from (2.1) or (2.4), depending on the way we enter the block effects into the model, $\boldsymbol{\beta} = (\beta_1,\ \beta_2,\ ...,\ \beta_{b-1})^t$, $\boldsymbol{\Sigma}_{ji} = \boldsymbol{\Sigma}_{ji}(\boldsymbol{\theta})$ as discussed in section 3, and $\boldsymbol{\theta} = (\theta_1,\ \theta_2, ..., \theta_G)^t$, the dispersion parameters.

Random blocks

When blocks are random, we no longer have independence for the subject-response vectors $\mathbf{y}_{ji}$. In this case, the independence occurs for the block-response vectors $\mathbf{y}_j$ and, following section 2 with the random block effects normally distributed, the log-likelihood function can be written as

$$\begin{aligned} L(\boldsymbol{\alpha}, \sigma_b^2, \boldsymbol{\theta}) &= (1/2)log(2\pi) \sum_{j=1}^{b} p_j - (1/2) \sum_{j=1}^{b} log|\boldsymbol{\Sigma}_j| \\ & -(1/2) \sum_{j=1}^{b} [\mathbf{y}_j - \mathbf{f}(\mathbf{X}_j, \boldsymbol{\alpha}_j)]^t \boldsymbol{\Sigma}_j^{-1} [\mathbf{y}_j - \mathbf{f}(\mathbf{X}_j, \boldsymbol{\alpha}_j)] \quad (4.2) \end{aligned}$$

with $\boldsymbol{\Sigma}_j$ given by (2.3) or (2.6), depending on the way we entered the block effect into the model, $\boldsymbol{\alpha}_j$ and $\boldsymbol{\Sigma}_j = \boldsymbol{\Sigma}_j(\boldsymbol{\theta})$ as above, and σ_b^2 as in (2.1).

The maximum likelihood (ML) estimators are the values of $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ in (4.1) and of $\boldsymbol{\alpha}$, σ_b^2 and $\boldsymbol{\theta}$ in (4.2) that maximize the corresponding function. There are no explicit solution for the likelihood equations which are obtained by equating the partial derivatives of (4.1) or (4.2), with respect to the corresponding parameters, to zero and numerical techniques such the Newton-Raphson or Fisher-Scoring algorithms are required. Restricted maximum likelihood (REML) estimators for the dispersion parameters $\boldsymbol{\theta}$ (and also for σ_b^2 for random blocks) can also be obtained by maximizing the likelihood of a linear transformation of the data for which the distribution does not depend on α (and also on β for fixed blocks), see Vonesh and Chinchilli (1997) for details. Notice that when the random block effects enter the model in a nonlinear fashion, the block-response vectors $\mathbf{y}_j$ would be only approximately normally distributed because of the first-order Taylor expansion considered.

The necessary modifications to use the existing methods for completely randomized designs in our randomized block setup, would be to:

1. add β_j, the effect of the $j-th$ block, to $\mathbf{f}(\mathbf{X}_j, \alpha_j)$ without affecting the covariance structure, for fixed blocks;

2. consider the nonlinear mixed-effects models setup with block effects as the random effects, for random blocks. In this case, there are only b independent response vectors, insted of $\sum_{j=1}^{b} n_j$ as in the completely randomized setup.

Confidence intervals and hypothesis testing can be done from asymptotic theory for (restricted) maximum likelihood. Nested models can be compared via likelihood ratio tests and the Akaike Information Criterion (AIC) can be considered to compare both nested and nonnested models. The reader is referred to Davidian and Giltinan (1995-Chapter 6) among others, for details. Additionally, Wald statistics can be considered to construct confidence intervals and hypothesis testing about the location parameters (see Sen and Singer (1993) for details), but they are not likely to be as reliable as inference based on likelihood principles.

5. Data Analysis

In this section we analyse the data presented in Table 5, whose parallel plots for the subjects are displayed in Figure 1. The main goal is not to fully analyse the data but to exemplify the methodology discussed above.

From section 1 and Figure 1 we can see that one appropriate function to model the growth of eucalyptus trees, measured by their volumes, would be a Gompertz function and that the

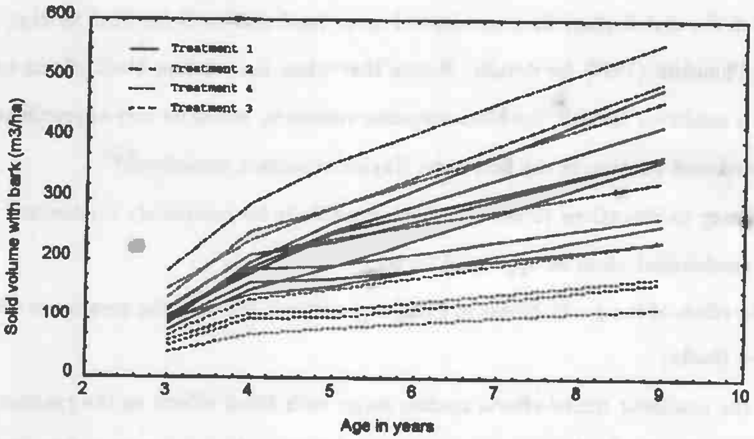

Figure 1: Parallel plots of solid volume with bark for the 16 subjects.

variability among the observations increase with time. Also, as in most of the practical situations, the conclusions we would like to draw should not be valid only for the blocks considered in the study. Therefore, we will treat the data as the 4 blocks were a random sample from a large population of blocks, i.e. we will treat the block effects as random. Also, due to the small sample size, we can not consider covariance structures with too many dispersion parameters.

Following our notation, we have model (2.1) or (2.5) with the elements of the nonlinear function $\mathbf{f}$ given by (1.2), for each one of the 4 treatments considered in the study, the covariance matrix $\boldsymbol{\Sigma}_j$ given by (2.3) or (2.6), $s = 1,\ 2,\ 3,\ 4$ (the instants of observation), $p_{ji} = 4$, $b = 4$, $n_j = 4$, $\mathbf{X}_{ji} = (x_{ji1}\ x_{ji2}\ x_{ji3}\ x_{ji4})^t = (3\ 4\ 5\ 9)^t$, $\mathbf{A}_j = \mathbf{I}_{12}$ and $\boldsymbol{\alpha}_j = \boldsymbol{\alpha} = (\boldsymbol{\alpha}_1\ \boldsymbol{\alpha}_2\ \boldsymbol{\alpha}_3\ \boldsymbol{\alpha}_4)^t$ with $\boldsymbol{\alpha}_k = (\alpha_{0k},\ \alpha_{1k},\ \alpha_{2k})^t$. Notice that, as the number of treatments is equal to the number of subject per block (complete randomized block design (CRBD)), there is an one-to-one relationship between i and k for each j. So, without loosing generality, we will replace k by i.

Two different structures for $\boldsymbol{\Sigma}_{ji}$, each one with only one parameter, are considered:

1. Independence: $\boldsymbol{\Sigma}_{ji} = \sigma^2 \mathbf{I}_4$

2. Simple dependence-heterocedasticity: $\boldsymbol{\Sigma}_{ji} = \sigma^2 \mathbf{R}$, with $\mathbf{R}$ given in (3.4).

The block effects are considered entering in the model in both linear (see model (2.1)) and nonlinear (see models (2.4) and (2.5)) fashions, with two different values for **b** in the nonlinear fashion:

(a) $\mathbf{b} = (1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1)^t$, generating

$$f_s(\mathbf{X}_{ji}, \boldsymbol{\alpha}_i) = (\alpha_{0i} + \beta_j) exp\{-exp[-(\alpha_{1i} + \beta_j)(x_{jis} - (\alpha_{2i} + \beta_j)/(\alpha_{1i} + \beta_j))]\};$$

(b) $\mathbf{b} = (1\ 0\ 0\ 1\ 0\ 0\ 1\ 0\ 0\ 1\ 0\ 0)^t$, generating

$$f_s(\mathbf{X}_{ji}, \boldsymbol{\alpha}_i) = (\alpha_{0i} + \beta_j) exp\{-exp[-\alpha_{1i}(x_{jis} - \alpha_{2i}/\alpha_{1i})]\}.$$

For this model, the $s-th$ element of $\mathbf{Z}_{ji}(\boldsymbol{\alpha}_i)$ would be

$$z_{jis}(\boldsymbol{\alpha}_i) = exp\{-exp[-\alpha_{1i}(x_{jis} - \alpha_{2i}/\alpha_{1i})]\}.$$

This last model assumes that the block effect interacts only with the asymptote parameters. A completely randomized design (CRD) setup was also considered to allow an evaluation of the importance of introducing block effects in the analysis. A total of 8 models were considered and the results are presented in Table 2. Comparisons between models 1 and 3, 2 and 4, 1 and 5, 2 and 6, 1 and 7, and 2 and 8 evaluate the block effects and can be done via likelihood ratio tests. By the other hand, AIC procedure can be utilized to compare any pair of these 8 models. Models 2 and 4 are the ones that presented the highest AIC's values. Comparing them via likelihood ratio test we have

$$-2log\,(\lambda_{REML}) = 506.74 - 488.59 = 18.15$$

which is much higher than 3.84, the tabled value of the chi-square with 1 degree of freedom and 5% significance level. Additionally, we can compare linear x nonlinear blocking by comparing the AIC's values for models 3 and 5, and 4 and 6. In both cases the AIC's values

Table 2: Results for the Akaike Information Criterion (AIC) and for -2log(l) for 8 different covariance structure models.

covariance model	design	structure for Σ_{ji}	AIC		-2log(l)[1]	
			ML	REML	ML	REML
1	CRD	1	-335.32	-327.76	644.65	629.52
2	CRD	2	-271.85	-266.37	517.70	506.74
3	CRBD with linear blocking	1	-314.35	-310.54	600.69	593.07
4	CRBD with linear blocking	2	-261.99	-258.29	495.99	488.59
5	CRBD with nonlinear blocking(a)	1	-319.43	-314.67	610.87	601.34
6	CRBD with nonlinear blocking(a)	2	-311.18	-307.68	594.36	587.36
7	CRBD with nonlinear blocking(b)	1	-298.24	-297.31	568.48	566.63
8	CRBD with nonlinear blocking(b)	2	-294.42	-293.95	560.84	559.90

[1]l: likelihood

associated to the linear blocking are higher. We select model 4 covariance structure as the most appropriate, among the 8 models considered, for our data set.

A second step of the analysis would be to make treatment comparisons assuming the covariance structure given by model 4 above. Table 3 presents the results for some of these comparisons. From the result obtained for the first hypothesis above, we can see that it is not appropriate to represent the mean volume as a function of age by only one curve for the four treatments. In the others 3 hypothesis, treatment effects are tested in terms of the asymptote parameters, and the conclusions are that, at a 5% significance level: there is no Species x Spacing interaction (hypothesis 2), there is Species main effect (hypothesis 3) and there is marginal Spacing main effect (hypothesis 4).

The last 2 hypothesis were also tested considering model 2 covariance structure (CRD) and the corresponding p-values were respectively 0.0697 and 0.1213, showing that the in-

Table 3: Likelihood ratio tests for treatment comparisons.

Null hypothesis	$-2log(l_{ML})$	$-2log(\lambda_{ML})$	d.f.	p-value
$\alpha_{01} = \alpha_{02} = \alpha_{03} = \alpha_{04}$ $\alpha_{11} = \alpha_{12} = \alpha_{13} = \alpha_{14}$ $\alpha_{21} = \alpha_{22} = \alpha_{23} = \alpha_{24}$	611.93	115.94	9	< 0.0000
$\alpha_{01} - \alpha_{02} = \alpha_{03} - \alpha_{04}$	497.75	1.76	1	0.1846
$\alpha_{01} + \alpha_{02} = \alpha_{03} + \alpha_{04}$	501.21	5.22	1	0.0223
$\alpha_{01} + \alpha_{03} = \alpha_{02} + \alpha_{04}$	499.81	3.82	1	0.0506

clusion of block effect into the model allowed us to detect Species and Spacing main effects. Figure 2 displays the final models for the four treatments, whose parameters estimates and their standard errors are presented in table 4.

All the calculations were made using specific computer routines developed by the authors utilizing procedures and the matrix language of the statistical package SAS. These calculations could also be done with the computer routines MIXNLIN (see Vonesh and Chinchilli (1997)) and the one cited in Pinheiro, Bates and Lindstrom (1993).

6. Conclusions

This work was mainly motivated by the fact that the methods referred in the literature to analyse nonlinear longitudinal date are suitable to data obtained from completely randomized designs. Two different ways of introducing block effects, in a linear or in a nonlinear fashion, were suggested for both fixed and random block effects. With random block effects, as it is in most of the practical applications, there are two types of dependence among the data: one due to the longitudinal character (among observations taken on the same subject) and one due to the random block effects (among observations taken on subjects belonging to the same block). So different dispersion parameters have to be considered to take it into account. By the other hand, in most of practical applications the number of blocks (and consequently

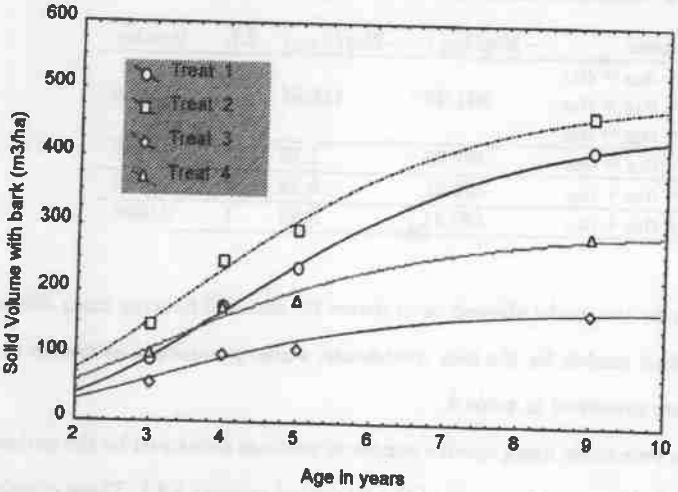

Figure 2: Estimated and observed means values for the four treatments.

Table 4: Maximum likelihood estimates of the parameters and their standard errors.

Parameters	Estimate	Standard error
α_1	461,28	36,43
α_2	503,54	33,28
α_3	180,15	27,06
α_4	303,45	30,42
β_1	1,76	0,24
β_2	1,49	0,20
β_3	1,58	0,59
β_4	1,37	0,34
γ_1	0,44	0,07
γ_2	0,44	0,06
γ_3	0,52	0,19
γ_4	0,46	0,11
σ_b^2	618,01	-[1]
σ^2	116,82	-

[1]values not calculated

the number of independent data) is not large. So, there is a need for considering covariance structure models for $\boldsymbol{\Sigma}_{ji}$ with very few parameters.

Despite the fact that our numerical example was based on balanced longitudinal data obtained from a complete randomized block design, one can see that our formulation is much more general, and allows the modelling of unbalanced longitudinal data and/or data obtained from incomplete randomized block designs. It also allows for a number of subject per block greater than the number of treatments. In this case, as there are more than one subject allocated to the same treatment in the same block, one can think of paralleling the nonlinear mixed-effects setup and introduce subject random effects into the model, and consequently into the modelling of the covariance matrix $\boldsymbol{\Sigma}_{ji}$.

Finally, the nonlinear way of introducing the block effects into the model allows different situations for introducing block x treatment and/or block x time interactions. For instance, in our example, one could imagine the block effects acting differently on each one of the 3 types of location parameters (α_0, α_1, α_2). In this case, it would be necessary to consider 3, instead of only one, different type of random block effects and at least three associated dispersion parameters (the block variance components). The existing methods for nonlinear mixed-effects models for longitudinal data obtained from completely randomized designs can be easily modified to the above situations.

Acknowledgments

This work was partially supported by grants from Coordenadoria para o Aperfeiçoamento de Pessoal de Ensino Superior (CAPES), Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA) under Technical Project $n^{\underline{o}}$ 23800.96/021-05-01, National Program of Excellency (PRONEX) contract $n^{\underline{o}}$ 76.97.1081.00 and Thematic Project of Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) convênio $n^{\underline{o}}$ 96/01741-7, and it was developed when the first author was a doctoral student at the Department of Mathematics and Statistics of Escola Superior de Agricultura Luiz de Queiroz - Universidade de São Paulo (ESALQ/USP), Piracicaba, São Paulo, Brazil. We wish to thank the Klabin Fabricadora de Papel e Celulose S.A. by borrowing the data set.

APPENDIX A: THE DATA SET FOR THE ANALYSIS

Table 5: Solid volume with bark (m^3/ha) of eucalyptus trees.

subject	block	species	spacing	treatment	age (years) 3	4	5	9
1	1	*E. grandis*	3.5 × 3.5	1	81.97	150.37	189.73	367.77
2	1	*E. grandis*	2.5 × 1.7	2	115.85	209.86	235.65	331.88
3	1	*E. dunnii*	3.5 × 3.5	3	43.83	74.78	85.78	121.83
4	1	*E. dunnii*	2.5 × 1.7	4	94.10	162.21	168.45	233.00
5	2	*E. grandis*	3.5 × 3.5	1	99.82	186.98	248.64	424.32
6	2	*E. grandis*	2.5 × 1.7	2	137.36	248.32	287.89	495.15
7	2	*E. dunnii*	3.5 × 3.5	3	63.60	108.02	115.58	170.27
8	2	*E. dunnii*	2.5 × 1.7	4	105.86	184.85	191.04	273.81
9	3	*E. grandis*	3.5 × 3.5	1	103.68	178.28	245.99	374.06
10	3	*E. grandis*	2.5 × 1.7	2	151.65	233.94	288.63	464.25
11	3	*E. dunnii*	3.5 × 3.5	3	57.01	101.03	102.54	162.07
12	3	*E. dunnii*	2.5 × 1.7	4	99.29	141.51	156.41	260.45
13	4	*E. grandis*	3.5 × 3.5	1	91.64	195.67	263.58	486.04
14	4	*E. grandis*	2.5 × 1.7	2	182.76	294.91	362.26	560.96
15	4	*E. dunnii*	3.5 × 3.5	3	72.47	134.27	150.43	233.21
16	4	*E. dunnii*	2.5 × 1.7	4	119.29	209.89	230.02	375.00

REFERENCES

Andrade,D.F., and Helms,R.W. (1986), "ML Estimation and LR Tests for the Multivariate Normal Distribution with General Linear Model Mean and Linear-Structure Covariance Matrix: K-Population, Complete-Data Case," *Communications in Statistics, Theory and Methods*, 15, 89-107.

Andrade,D.F., and Singer,J.M. (1998), "Profile Analysis for Randomized Complete Block Experiments," *Journal of Applied Statistics*, 25, 237-244.

Crowder,M.J., and Hand, D.J. (1995), *Analysis of Repeated Measures*, London: Chapman & Hall.

Davidian,M., and Giltinan, D.M. (1995), *Nonlinear Models for Repeated Measurement Data*, London: Chapman and Hall.

Gallant,A.R. (1987), *Nonlinear Statistical Models*, New York: Wiley.

Gennings,C., Chinchilli,V.M., and Carter,W.H. (1989), "Response Surface Analysis with Correlated Data: A Nonlinear Model Approach," *Journal of the American Statistical Association*, 84, 805-809.

Goldstein,H. (1979), *The design and analysis of longitudinal studies*, New York: Academic Press.

Lindsey,J.K. (1993), *Models for the repeated measurements*, Oxford: Oxford University Press.

Lindstrom,M.J., and Bates,D.M. (1990), "Nonlinear Mixed Effects Models for Repeated Measures Data," *Biometrics*, 46, 673-687.

Neter,J., Wasserman,W., and Kutner, M.H. (1990), *Applied Linear Statistical Models: Regression, Analysis of Variance, and Experimental Designs*, Boston: Irwin, Inc.

Nokoe,S. (1980), "Nonlinear Models Fitted to Stand Volume-Age Data Compare Favourably with British Columbia Forest Service Hand-Drawn Volume-Age Curves," *Canadian Journal of Foresty Research*, 10, 304-307.

Pinheiro,J.C., Bates,D.M., and Lindstrom,M.J. (1993), "Nonlinear Mixed Effects Classes and Methods for S," *Technical Report nº 906*, University of Wisconsin-Madison, Dept. of Statistics..

Ratkowsky,D.A. (1983), *Nonlinear Regression Modeling - An Unified Pratical Approach*, New York: Marcel Dekker.

Seber,G.A.F., and Wild, C.J. (1989), *Nonlinear Regression*, New York: John Wiley.

Sen,P.K., and Singer, J.M. (1993), *Large Sample Methods in Statistics: An Introduction with Applications*, New York: Chapman and Hall.

Singer,J.M., and Andrade,D.F. (1994), "On the Choice of Appropriate Error Terms in Profile Analysis," *The Statistician*, 43, 251-261.

Vonesh,E.F. (1992), "Non-Linear Models for the Analysis of Longitudinal Data," *Statistics in Medicine*, 11, 1929-1954.

Vonesh,E.F., and Carter,R.L. (1992), "Mixed-Effects Nonlinear Regression for Unbalanced Repeated Measures," *Biometrics*, 48, 1-17.

Vonesh,E.F., and Chinchilli, V.M. (1997), *Linear and Nonlinear Models for the Analysis of Repeated Measurements*, New York: Marcel Dekker, Inc.

Wolfinger,R.D. (1996), "Heterogeneous Variance-Covariance Structures for Repeated Measures," *Journal of Agricultural, Biological, and Environmental Statistics*, 1, 205-230.

ÚLTIMOS RELATÓRIOS TÉCNICOS PUBLICADOS

9801 - **GUIOL, H.** Some properties of K-step exclusion processes. 1998. 21p. (RT-MAE-9801)

9802 - **ARTES, R., JØRSENSEN, B.** Longitudinal data estimating equations for dispersion models. 1998. 18p. (RT-MAE-9802)

9803 - **BORGES, W.S.; HO, L.L.** On capability index for non-nornal processes. 1998. 13p. (RT-MAE-9803)

9804 - **BRANCO, M.D. , BOLFARINE, H., IGLESIAS, P., ARELLANO-VALLE, R.** Bayesian analysis of the calibration problem under elliptical distributions. 1998. 15p. (RT-MAE-9804)

9805 - **IGLESIAS, P., MATÚS, F., PEREIRA, C.A.B., TANAKA, N.I.** On finite sequences conditionally uniform given minima and maxima. 1998. 13p. (RT-MAE-9805)

9806 - **IGLESIAS, P., PEREIRA, C.A.B., TANAKA, N.I.** Characterizations of multivariate spherical distribution in l_∞ norm. 1998. 19p. (RT-MAE-9806)

9807 - **GIAMPAOLI, V., SINGER, J.M.** A note on the comparison of two normal populations with restricted means. 1998. 12p. (RT-MAE-9807)

9808 - **BUENO, V.C.** A note on conditional reliability importance of components. 1998. 07p. (RT-MAE-9808)

9809 - **GIMENEZ, P., BOLFARINE, H.** Consistent estimation in comparative calibration models. 1998. 19p. (RT-MAE-9809)

9810 - **ARTES, R., PAULA, G.A., RANVAUD, R.** Estudo da direção tomada por pombos através de equações de estimação para dados circulares longitudinais. 1998. 16p. (RT-MAE-9810)

9811 - **LIMA, A.C.P., SEM, P.K.** Time-Dependent Coefficients in a Multi-Event Model for Survival Analysis. 1998. 21p. (RT-MAE-9811)

9812 - **KOLEV, N., MINKOVA, L.** Over-and underdispersed models for ruin probabilities. 1998. 34p. (RT-MAE-9812)

9813 - **KOLEV, N.** Negative Binomial Cross-Classifications. 1998. 16p. (RT-MAE-9813)

9814- **BAKEVA, V., GEORGIEVA, M., KOLEV, N.** Two Characterizations of the Geometric Distribution Related to an Unreliable Geo/G_D/1 System. 1998. 11p. (RT-MAE-9814)

9815 - **KOLEV, N., MINKOVA, L.** Some basic Inflated-Parameter Discrete Distributions. 1998. 19p. (RT-MAE-9815)

9816 - **ARTES, R., PAULA, G.A., RANVAUD, R.** Strategy for the analysis of circular longitudinal data based on generalized estimating equations. 1998. 15p. (RT-MAE-9816)

9817 - **PAULA, G.A., ARTES, R.** A Score-Type Test to Assess Overdispersion in Linear Logistic Models. 1998. 17p. (RT-MAE-9817)

9818 - **LOSCHI, R.H., IGLESIAS, P.L., WECHSLER, S.** Unpredictability and Probability Updating. 1998. 11p. (RT-MAE-9818)

9819 - **KOLEV, N., MINKOVA, L.** New Over-/Under- Dispersed Class of Inflated-Parameter Discrete Probability Distributions. 1998. 10p. (RT-MAE-9819)

9820 - **KOLEV, N., MINKOVA, L.** Inflated-Parameter Family of Generalized Power Series Distributions. 1998. 11p. (RT-MAE-9820)

9821 - **AUBIN, E.C.Q., STANGENHAUS, G.** A Robust Estimation of Growth Curves. 1998. 7p.(RT-MAE-9821)

9822 - **BORGES, W., HO, L. L.,** A Fraction Non-Conforming Based Capability Index. 1998. 13p. (RT-MAE-9822)

9823 - **GALEA-ROJAS, M., BOLFARINE, H.** Hypothesis Testing in Functional Comparative Calibration Models. 1998. 13p. (RT-MAE-9823)

9824 - **FONTES, L. R. G., ISOPI, M., NEWMAN, C. M.** Chaotic Time Dependence in a Disordered Spin System. 1998. 23p. (RT-MAE-9824)

9825 - **FERRARI, P. A., FONTES, L. R. G., VARES, M. E.** The Asymmetric Simple Exclusion Model With Multiple Shocks. 1998. 24p. (RT-MAE-9825)

9826 - **CABRAL, C. R. B., PAULA, G. A., BOLFARINE, H.** One-Sided Tests for the Analysis of Pretest-Posttest Data. 1998. 12p. (RT-MAE-9826)

9827 - **GASCO, L., BOLFARINE, H., ARELLANO-VALLE, R. B.** Asymptotic properties of the sample mean vector and sample covariance matrix with application in elliptical multiplicative measurement error models. 1998. 17p. (RT-MAE-9827)

9828 - **FERRARI, S. L. P., CRIBARI-NETO, F.** On the robustness of analytical and bootstrap corrections to score tests in regression models. 1998. 11p. (RT-MAE-9828)

9829 - **PERES, C. A., ANDRE, C. D. S., NARULA, S. C.** Whyn the observations within a treatment of a random-effects model are correlated? 1998. 8p. (RT-MAE-9829)

9830 - **FONTES, L.R., MATHIEU, P., PICCO, P.** On the averaged dynamics of the random field curie-weiss model. 1998. 27p. (RT-MAE-9830)

9831 - **ANDRÉ, C.D.S., ELIAN, S.N., TAVARES, R.** Coeficientes de determinação na regressão de mínima soma dos erros absolutos e sua utilização em problemas de seleção de variáveis. 1998. 14p. (RT-MAE-9831)

9832 - **SOUZA, F.A.M., PAULA, G.A.** Deviance residuals for an angular response. 1998. 13p. (RT-MAE-9832)

9833 - **SOLER, J.M.P., SINGER, J.M.** Optimal covariance adjustment in growth curve models. 1998. 13p. (RT-MAE-9833)

9834 - **BOTTER, D.A., PERES, A.A.Q., ZANDONADE, E.** A longitudinal study of mortality and air pollution for São Paulo, Brazil. 1998. 16p. (RT-MAE-9834)

9835 - **CRIBARI-NETO, F., FERRARI, S.L.P.** Monotonic improved critical values for econometric chi-squared asymptotic criteria. 1998. 12p. (RT-MAE-9835)

9836 - **CORDEIRO, G.M., FERRARI, S.L.P., URIBE-OPAZO, M.A., VASCONCELLOS, K.L.P.** Corrected Maximum Likelihood estimation in a class of symmetric nonlinear regression models[1]. 1998. 14p. (RT-MAE-9836)

9837 - **BELITSKY, V., FERRARI, P.A.** Invariant measures and convergence for cellular automaton 184 and related processes. 1998. 19p. (RT-MAE-9837)

9838 - **CASSANDRO, M., COLLET, P., GALVES, A., GALVES, CHARLOTTE.** A Statistical-Physics Aprroach to Language Acquisition and Language Change. 1998. 15p. (RT-MAE-9838)

9839 - ANDRADE, D.F., VALLE, R.C. Uma introdução à teoria da resposta ao item : conceitos e aplicações. 1998. 15p. (RT-MAE-9839)

9840 - BUENO, V.C. Component importance in a random environment. 1998. 6p. (RT-MAE-9840)

The complete list of "Relatórios do Departamento de Estatística", IME-USP, will be sent upon request.

Departamento de Estatística
IME-USP
Caixa Postal 66.281
05315-970 - São Paulo, Brasil