Formulating mixed models for experiments, including longitudinal experiments

C. J. Brien[1] and C.G.B. Demétrio[2]

[1]School of Mathematics and Statistics, University of South Australia,

North Terrace, Adelaide, SA 5000, Australia.

[2]Departmento de Ciências Exatas, ESALQ, Universidade de São Paulo,

Piracicaba, SP 13418-900, Brasil.

August 17, 2009

Mixed models have become important in analyzing the results of experiments, particularly those that require more complicated models (e.g. those that involve longitudinal data). This article describes a method for deriving the terms in a mixed model. Our approach extends an earlier method by Brien and Bailey to explicitly identify terms for which autocorrelation and smooth trend arising from longitudinal observations need to be incorporated in the model. At the same time we retain the principle that the model used should include, at least, all the terms that are justified by the randomization. This is done by dividing the factors into sets, called tiers, based on the randomization and determining the crossing and nesting relationships between factors. The method is applied to formulate mixed models for a wide range of examples. We also describe the mixed model analysis of data from a three-phase experiment to investigate the effect of time of refinement on Eucalyptus pulp from four different sources. Cubic smoothing splines are used to describe differences in the trend over time and unstructured covariance matrices between times are found to be necessary.

Keywords: Analysis of variance; Longitudinal experiments; Mixed models; Multiphase experiments; Multitiered experiments; Repeated measures.

1. Introduction

As Piepho, Büchse, and Emrich (2003) (henceforth P3) noted, there are many situations in which the use of linear mixed model software, rather than ANOVA software, is preferred. These include cases where nonorthogonality is involved and where models more complicated than the homogeneous ANOVA-models are required. Piepho, Büchse, and Richter (2004) (henceforth P4) imparted a method for deriving such mixed models for experiments involving longitudinal observations, the models assuming unequal correlations and variances between observations at different times on the same unit. Brien and Bailey (2006) also described a method for formulating the mixed model for an experiment without explicitly describing how data with longitudinal observations might be handled. These methods, along with that for deriving an ANOVA table given by Brien and Demétrio (1998), are based on Nelder (1965a, b) and Brien (1983) and emphasize the principle of basing a mixed model on the randomization employed in the experiment. Other authors who advocated this principle, by also initially considering the topographical, physical or managerial features of an experiment separately from the treatment features, include Fisher (1935), Wilk and Kempthorne (1956), White (1975) and Bailey (1981). Using a different mechanism, Lorenzen and Anderson (1993) also emphasized this principle. However, the principle is often compromised, such as when factors indexing the units are omitted and randomized treatment factors substituted for them. For example, in an experiment employing a randomized complete block design (RCBD), the Plots factor is rarely included in the model; instead the Treatments factor acts as a surrogate for it. This practice is encapsulated in Rule 5 of P3.

As in P4, it is important to be clear that by *longitudinal experiments* we mean those in which successive observations have been made on *some* unit over time or space and without randomization to these times or positions. We differ from P4 in that we do not insist that times or positions are unrandomized, only that factors are not randomized to them. Such experiments are a subset of *repeated measures experiments* as the latter term includes experiments in which factors are assigned to times, such as carry-over experiments.

The purpose of this article is to extend the method of Brien and Bailey (2006) to explicitly cover longitudinal experiments, as suggested by Piepho (2006), while simplifying the specification of mixed models for such experiments. The proposed method, like those from which it is derived, has the feature that the model for an experiment is derived from first principles, rather than by using the model for the experiment that most closely matches it in a list of standard designs. Such an approach has been labeled "no-name" by Lorenzen and Anderson (1993), as the name of the design is not needed to determine its model. Section 2 gives a motivating example, and Section 3 outlines the notation to be used for mixed models. Section 4 describes the proposed method, and Section 5 applies the method to the P4 examples. Section 6 describes the extension to multitiered experiments, and Section 7 analyzes the data from the three-phase experiment introduced in Section 2. Section 8 discusses the differences between our method and that of P4 and the problems that arise from not using randomization-based models. Section 8 also discusses problems that arise from inextricable confounding of terms, especially when there is improper replication or pseudoreplication (Hurlbert, 1984; Brien and Demétrio, 1998).

2. A motivating example

Pereira (1969) described an experiment, that is in fact three phase, to investigate differences between pulp produced from different eucalypt trees. First, in the chip phase, three lots of five trees from each of four areas, that differed in kinds (species *E. saligna* Smith and *E. alba* Reinw) and age (5 and 7 years) of trees, were selected and processed into wood chips. (A third species, *E. grandis*, was omitted from the analysis because the selection of trees was much more limited for it.) For each of the 12 lots, the chips from five trees were combined and four batches were selected. Second, in the pulp phase, batches were cooked to produce pulp and six samples obtained from each cooking. Third, in the measurement phase, each batch was processed in one of 48 Runs of a laboratory refiner with its six samples randomly placed on six positions in a pan in the refiner. For each run, six times of refinement (30, 60, 90, 120, 150 and 180 min) were randomized to the six positions in the pan. At the end of its allotted time of refinement a sample was taken from the pan and its degree of refinement (°SR) measured using the Schopper–Riegler method. Ideally, there would be randomization of Batches to Cookings and of Cookings to Runs, as illustrated in the randomization diagram (Brien and Bailey, 2006) in Figure 1. This diagram encapsulates the experimental units for this experiment and the restrictions placed on randomization. Both Runs and Positions (in Runs) are experimental units, as they had factors randomized to them, and the randomization was restricted to positions within runs as indicated by the nesting.

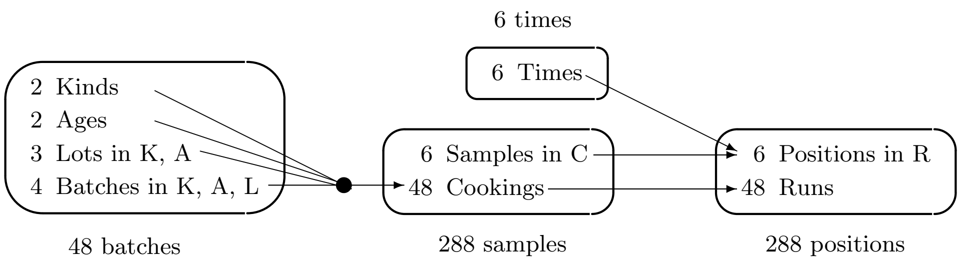

Figure 1. Potential randomizations for the Eucalypt pulp experiment

Appendix A contains the results in systematic order. They are displayed in Figure 2, from which it is evident that there is curvature in the trend over time and variance heterogeneity, in particular between the Ages. A mixed-model analysis of this example will be described in Section 7.

3. Mixed models and symbolic notation

A general form for the linear mixed model is (Robinson, 1991; Wolfinger, 1993):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

where $\mathbf{y}$ is an n-vector of observations, $\boldsymbol{\beta}$ is the p-vector of parameters for fixed effects, $\mathbf{X}$ is an $n \times p$ indicator-variable matrix for fixed effects, $\mathbf{Z}$ is an $n \times q$ indicator-variable matrix for random effects, and

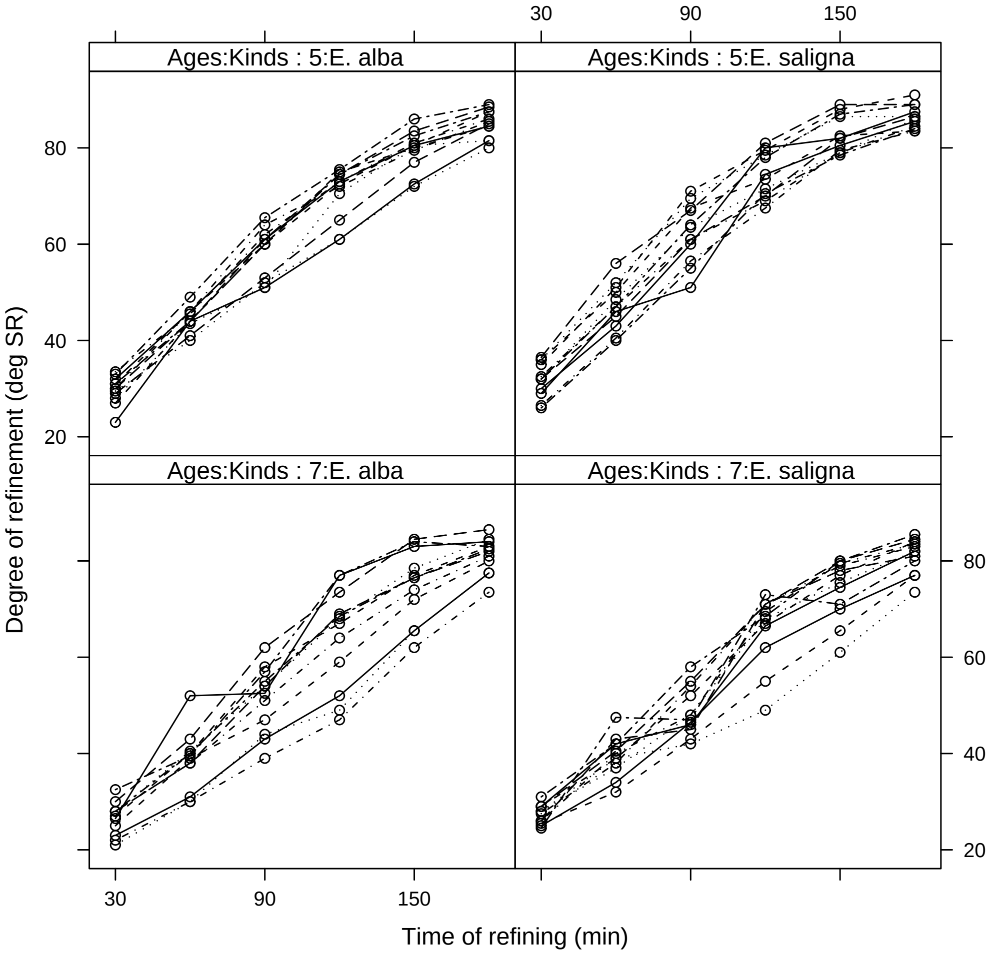

Figure 2. Degree of refinement for each batch in the Eucalypt pulp experiment

u and **e** are the q- and n-vectors, respectively, of unobservable random effects for which it is assumed that

$$\mathrm{E}\left[\mathbf{u}\right] = \mathbf{0}, \mathrm{E}\left[\mathbf{e}\right] = \mathbf{0} \text{ and } \mathrm{Var}\begin{bmatrix} \mathbf{u} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{G} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix}$$

where **G** and **R** are positive definite matrices. The multilevel models of Laird and Ware (1982) and Pinheiro and Bates (2000) can be expressed in this form. It is sometimes called the conditional model and it implies the marginal model: $E[\mathbf{y}] = \mathbf{X}\boldsymbol{\beta}$ and $\mathrm{Var}[\mathbf{y}] = \mathbf{V} = \mathbf{ZGZ}' + \mathbf{R}$ (Littel et al., 2006). However, a marginal model does not necessarily imply a conditional model, such as when negative variance components are allowed. The fixed part of the model can be modified to include nonlinear models.

The objective of the proposed method is to derive, in symbolic form, the fixed model that consists of the model terms corresponding to the submatrices of **X** and the random model with model terms corresponding to the submatrices of **Z**, as well as a unit term corresponding to **R**. For specifying the model we employ a similar convention to Patterson (1997) and P3. The terms for the fixed model are given first, followed by the terms for the random model. The two models are separated by a "|" as in R (R Core Development Team, 2008), rather than a colon as in P3. The notation to be used in specifying each of these models is summarized in Table 1. The *model terms* are expressed in terms of generalized factors (Brien and Bailey, 2006), where a *generalized factor* is a list of factors separated by the "wedge" operator ($\wedge$) and is the new factor whose levels are the observed combinations of the levels of its constituent factors. This is the same as in P3, except that the "·" is replaced by the " $\wedge$". If A has a levels and B has b levels and all combinations of A and B are observed, then A $\wedge$B has ab levels. Corresponding to A $\wedge$B is the $n \times ab$ submatrix $\mathbf{X}_{\mathrm{A}\wedge\mathrm{B}}$. A *unit term* is one whose factors uniquely index the observational units; if there are more than one, then it will be necessary to choose one to correspond to **R**.

Generally, for compactness, we will not list out all the terms in a model but give them as *model*

Table 1. Table summarizing mixed model notation

Wedge operator ($\wedge$)	
(i)	A $\wedge$B = B $\wedge$A is the generalized factor formed from the observed combinations of the levels of A and B.
(ii)	(A $\wedge$B) $\wedge$C = A $\wedge$(B $\wedge$C)
(iii)	(A $\wedge$C) $\wedge$(B $\wedge$C) = A $\wedge$B $\wedge$C
(iv)	$L \wedge M$, where L and M are model formulas, is expanded by taking all pairwise products of the expanded elements of L with those of M.
Additive operator (+)	
(i)	(A+B) $\wedge$(C+D) = A $\wedge$C+A $\wedge$D+B $\wedge$C+B $\wedge$D
Generalized factor operator [gf(.)]	
(i)	forms the generalized factor from all the factors appearing in the model that is its argument.
(ii)	gf(L) = A $\wedge$B $\wedge$C . . . where A, B, C . . . are factors in L.
(iii)	(A + B)/C = A + B + gf(A + B) $\wedge$ C = A + B + A $\wedge$ B $\wedge$ C
Nesting operator (/)	
(i)	$L \,/\, M = L + \text{gf}(L) \wedge M$
(ii)	$L \,/\, (M \,/\, N) = (L \,/\, M) \,/\, N$
(iii)	$L \,/\, (M + N) = (L \,/\, M) + (L \,/\, N)$
Crossing operator ($*$)	
(i)	$L * M = L + M + L \wedge M$
(ii)	$L * (M + N) = (L * M) + (L * N)$
(iii)	$L * M \,/\, N = L * (M \,/\, N)$
Trend operators	
td(.)	specifies that models for trend in the response, as a function of the values of the quantitative variable corresponding to the factor, are to be investigated.
lin(.)	specifies a linear trend on quantitative factors.
pol(.)	specifies a polynomial trend on quantitative factors.
nln(.)	specifies nonlinear trend on quantitative factors.
spl(.)	specifies that overall trends are to be modeled using cubic smoothing splines by fitting a variance component to the random model.
dev(.)	specifies random deviations around the fitted trend.
Correlation matrix operators	
uc(.)	Unequal correlation operator specifying that, in the random model, unequal, possibly structured, correlations between different levels of the (generalized) factor will be considered.
ar1(.)	Autoregressive correlation of order 1 operator.
corb(.)	Banded correlation matrix operator.
us(.)	Unrestricted correlation matrix operator.
Unequal variances	
	add an "h" to the name of a correlation matrix operator.

formulas using factor relationship operators of nesting, crossing, and additive, similar to that used in P3 but restricted to standard ASCII symbolic characters as shown in Table 1. Such formulas are expanded to generalized factors or model terms using the rules given in Table 1. The "gf" operator, used in this, is more general than the product term operator in P3 in that it combines factors even when all combinations are not observed.

Two general functions, the trend and unequal correlation operators, are described in Table 1 and will be applied to (generalized) factors. The *trend operator* [td(.)] indicates that consideration is to be given to the form of the $\mathbf{X}$ matrix, at least, while *unequal correlation operator* [uc(.)] indicates that alternative forms of the **G** and **R** matrices are to be considered. These two general operators will be sufficient to indicate, prior to receiving the data, the approach to be taken in analyzing the results of an experiment. However, in performing the analysis it will be desirable to indicate precisely how the trend and correlation are being modeled, by replacing the general operators with specific ones such as in Butler et al. (2007) and Verbyla et al. (1999). A number of them are listed in Table 1. They are applied to single (generalized) factors.

The trend could be modeled in a number of ways. For example, lin(.) in the fixed model could be used to indicate linear trends over the levels of quantitative factors, while in the random model spl(.) could indicate that overall trends are to be modeled using cubic smoothing splines and dev(.) that random deviations around the fitted trend are to be included in the model. To allow the trends to vary randomly, as in random coefficients regression, any of the trend operators would be combined with subject terms in the random model (Section 4).

Similarly, the unequal correlation operator could be replaced by particular functions to specify the form of the correlation. For example, ar1(.) would indicate autoregressive correlation of order 1 associated with the generalized factor and ush(.) would indicate an unstructured covariance matrix. See Section 7 for an example of the use of these specific operators.

Having obtained the symbolic fixed and random models they will need to be supplied, using the appropriate syntax, to mixed model software such as that implemented in the ASReml add-ins for S-PLUS and R (Butler et al., 2007), GenStat (Payne et al., 2008), R (R Core Development Team, 2008), SAS (Littel et al., 2006) and S-PLUS (Insightful Corporation, 2005).

4. Method for two-tiered experiments

Many experiments are two-tiered, in that there is a single randomization in which one *set of objects*, usually called the *treatments*, is randomized onto a second set of objects, the *units*. There will be a *tier*, or set of factors, indexing these two sets of objects: the *unrandomized factors* index the units and the *randomized factors* index the treatments (Brien and Bailey, 2006). This is equivalent to considering the factors involved in block and treatment structures as proposed by Nelder (1965a, b) or the experiment and treatment structures as suggested by Littel et al. (2006, Section 4.2).

In describing the method, Example 1 of P4 will be used as an illustration. It involves a field experiment comparing three different tillage methods (A) and is laid out according to an RCBD with four blocks. In each plot there are four layers and one water collector is installed per layer to measure the amount of nitrogen leaching.

The proposed method is a three-stage process, outlined below and summarized in Figure 3:

1. Determine formulas for the intratier model made up of an intratier random (IR) model and an intratier fixed (IF) model:

   (a) Identify the sets of objects and nominate the observational unit. For two-tiered experiments there will be two sets of objects: the unrandomized and randomized sets. The *observational unit* (or unrandomized object) is the smallest unit from which a single value of a response variable is obtained. For longitudinal experiments, the observational unit is often a mixture of experimental and longitudinal units: a plot or subject at a time, a plot at a depth.

   Example 1: the sets of objects are i) 48 layer-plot combinations and ii) three treatments; the observational unit is a layer-plot combination.

   (b) Determine the two *tiers*: the unrandomized and randomized tier, each tier being the set of factors indexing the corresponding set of objects. A tier is comprised of the factors in the panel of a randomization diagram (Brien and Bailey, 2006).

   Example 1: The unrandomized tier is {Block, Plot, Lay} and the randomized tier is {A}.

   (c) Identify longitudinal factors, if any. *Longitudinal factors* are those to which no factors are randomized and that index successive observations of some unit.

   Example 1: The only longitudinal factor is Lay.

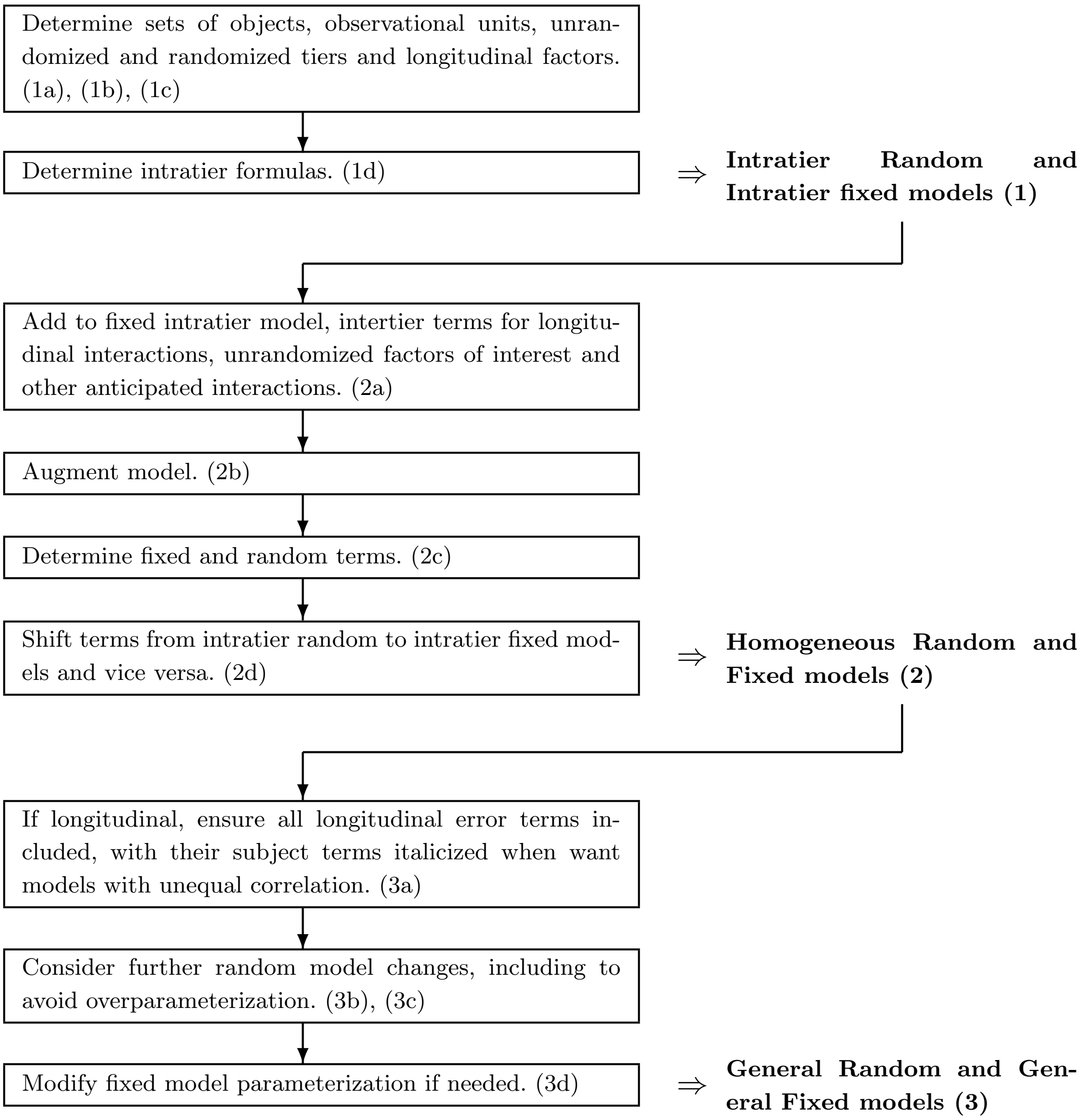

Figure 3. Three-stage process for formulating a linear mixed model

(d) Determine the *intratier formulas* by specifying the crossing, nesting and additive relationships between the factors in each tier. These relationships will depend on both the intrinsic relationships and those for the design employed (Brien and Bailey, 2006, Section 2.2). Usually longitudinal factors are crossed with other factors within a tier and, when possible, should be placed to the right of others; this is not possible for factors nested within longitudinal factors. The formula derived from the unrandomized factors specifies the intratier random model under randomization, or IR model, and that from the randomized factors specifies the intratier fixed model, or IF model. The implied marginal model is equivalent to the randomization model, except that the latter allows negative components of variance.

Example 1: IR: (Block/Plot)∗Lay; IF: A.

2. Convert the intratier model to a mixed model made up of a homogeneous random (HR) model and a fixed (F) model:

(a) Add to the IF model *intertier terms* that are of interest to the investigator (e.g. unrandomized Sites or Sex and their interactions with randomized factors) or thought to be important because they are likely to be appreciable (e.g., Judge interactions in sensory evaluation experiments). If there are longitudinal factors, then these will usually interact with randomized factors (treatments) and unrandomized factors of interest to the experimenter; such interactions that are not already included should also be added to the IF model.

Example 1: Differences for each level of Lay are of interest and the fixed model becomes F: A∗Lay.

(b) Augment the IR and IF models with potentially important terms not taken into account in the experimental design. These may involve factors that are identified as important sources of variation only after the experiment has begun or continuous covariates that could not be controlled for in the design. This is to be avoided if at all possible as it is preferable to account for all important sources of variability at the design stage. For example, a block term might be added in the analysis of a completely randomized design when, during the conduct of the experiment, it is realized that there are differences between groups of units. Of course, Treatments and Blocks will be nonorthogonal and the efficiency of Treatments within Blocks may be poor.

Example 1: not required.

(c) Decide which terms are fixed and which are random. This can be done by deciding which factors are fixed and which are random and then using the rule that a fixed term involves only fixed factors, otherwise the term is random. Alternatively, one can decide on a term by term basis, although it is unlikely that a term will be fixed if a subset of its factors corresponds to a random term. This means that A ∧B will be random if A is random. However, a term for which all of its factor subsets are fixed may be random. Take a Latin square with Rows and Columns that are fixed—even so Rows ∧Columns must be random.

Example 1: The fixed factors are Block, A and Lay and the random factor is Plot.

(d) Obtain the final HR and F models by shifting a) terms that are fixed from the previously modified IR model to the final F model and b) terms that are random from the previously modified IF model to the final HR model.

Example 1: The IR and F terms are:

IR: Block + Block ∧Plot + Lay + Block ∧Lay + Block ∧Plot ∧Lay;
F: A + Lay + A ∧Lay.

After the interchange of terms, the models become:

HR: Block ∧Plot + Block ∧Plot ∧Lay = (Block ∧Plot)/Lay;
F: Block + Lay + Block ∧Lay + A + A ∧Lay = (Block+A)∗Lay.

3. Formulate the general mixed model by deriving the general random (GR) and general fixed (GF) models from the HR and F models:

(a) If there are longitudinal factors, allow for unequal correlation between their levels by ensuring that all longitudinal error terms have been included in the random model. *Longitudinal error terms* are random terms of the form "(subject term) ∧ gf(longitudinal factors)". A *subject term* for a longitudinal factor is a generalized factor whose levels are units on which the successive observations are taken; it is usually comprised of unrandomized factors that are not longitudinal, but on occasion will include randomized factors (Examples 5 and 8 in Section 5) and other longitudinal factors. The subject terms are italicized if models involving unequal correlation are to be considered for their longitudinal error terms. Usually a subject term itself will also be in the random model, while terms for the longitudinal factors on their own are in the fixed model. The "gf" operator is applied to the part of the formula involving longitudinal factors so that no restriction is placed on the form of the correlation between the levels combinations of the longitudinal factors. Then this generalized factor has the "uc" operator applied to it. If a longitudinal factor happens to be randomized, it will be necessary to incorporate its longitudinal error terms by adding interactions between it and its subject terms. A longitudinal factor may have several subject terms. If, as discussed in P4, unequal correlation models are to be restricted to terms involving only unrandomized factors, longitudinal error terms should be restricted to terms comprised only of such factors.

Example 1: The subject term is Block ∧Plot and it is expected that there will be unequal correlation between observations with different levels of Lay. Hence, the initial general random model is formed as follows:

$$\text{GR:} \quad (\text{Block} \wedge \text{Plot})/\,\text{gf}(\text{Lay}) \Rightarrow (\textit{Block} \wedge \textit{Plot})/\,\text{uc}(\text{Lay}).$$

(b) Make other changes to the GR model. For example, allow for unequal variances between levels of longitudinal factors or unequal correlations between levels of other random factors.

Example 1: not required.

(c) Consider whether any random terms need to be removed to avoid overparameterization.

Example 1: not required.

(d) If appropriate apply the "td" operator to specify other parameterizations for the fixed terms such as polynomial, smooth or nonlinear trends. In many instances systematic trends across the levels of longitudinal factors are of interest.

Example 1: Trends for Lay are of interest, but not for the qualitative factor A nor for Block.

$$\text{GF:} \quad (\text{Block+A}) * \text{td}(\text{Lay}).$$

It is often useful to have constructed a decomposition (or ANOVA) table (Brien and Bailey, 2009), essentially an analysis of variance table without data, for the model determined by the end of Step 2(a), as an aid to understanding the confounding that has occurred in the experiment (Example 10 in Section 5 and Section 7). It is especially helpful in predicting and diagnosing overparameterization that arises from inextricable confounding (Section 8). It can be obtained in GenStat (Payne et al., 2008), S-PLUS (Insightful Corporation, 2005), and R (R Core Development Team, 2008). In GenStat, the IR model is supplied in the `BLOCKSTRUCTURE` directive, the IF model with extra terms in the `TREATMENTSTRUCTURE` directive and an `ANOVA` directive without a response variable is issued to produce the table. In S-PLUS and R, a call to the `aov` function of the form `y ~ IF model + extra terms + Error(IR model)`, where `y` is a vector of data or randomly generated numbers, will produce the table.

When there are no longitudinal factors, so that Steps 1(c), 3(a) and parts of Step 2(a) can be omitted, this procedure reduces to that of Brien and Bailey (2006). As in P4, when there are longitudinal factors, a number of correlation structures (for example Wolfinger, 1993; Pinheiro and Bates, 2000; Verbeke and Molenberghs, 2000; Butler et al., 2007) as well as unequal variances should be considered; serial correlation is a prime contender for modeling correlation between the longitudinal points. However, our approach differs from P4 in that all subject terms are allowed to occur on their own which, as they pointed out, can be allowed for by suitably parameterizing longitudinal error terms. One benefit of this is that it makes the handling of randomized and unrandomized terms, that are to be assumed random, more alike; the only difference is whether they can be considered for subject terms. Another benefit in the context of the more general approach here, where longitudinal factors can be randomized, is that it is simpler to include explicitly separate subject terms. Furthermore, it ensures models that include such terms are considered. Thus we take the approach of attempting to fit the more complex random model first, although not the maximal model which is an unstructured $\mathbf{V}$, and then consider whether this can be simplified. Of course, it may be that there is insufficient data to allow these models to be fitted and reduced models will have to be considered. If the terms are found ultimately to be unnecessary, then their removal means that we revert to the same models as considered in P4. For Example 1, if Block ∧Plot is not significant, its removal from the GR models results in the model $\textit{Block} \wedge \textit{Plot} \wedge \text{uc}(\text{Lay})$, the same as in P4.

To fit the model an approach along the lines of Verbeke and Molenberghs (2000) would be used:

(i) **Baseline mixed model:** Fit the GR and GF models, ignoring the "uc" and "td" operators. This will fit the maximal model to be considered for the fixed effects.

(ii) **Random trend modeling:** If trends are to be allowed to vary randomly between the levels of one or more random factors, then a) the terms in the GF model involving these trends should be reparameterized to indicate the specific trend operator that is to be investigated and corresponding random trend deviation terms incorporated into the random model, along with spline terms if these are required; b) randomly varying trend terms, being the interactions between the trend terms and random factors, are added to the random model. In doing this, note that random intercepts that vary for a particular combination of factors are equivalent to a variance component for that set of factors and the inclusion of the random intercepts, usually as part of the "lin" operator, will require the omission of any corresponding variance component. In longitudinal experiments fixed trend terms are likely to involve the longitudinal factors and the randomly varying trend terms will be related to longitudinal error terms. Now fit the model and investigate which terms can be removed from the random model, starting with the highest order terms.

(iii) **Covariance structure:** Investigate alternative parameterizations and omissions of random terms in the original GR model.

(iv) **Fixed model:** Determine the significance of terms in the fixed model. If no random trend modeling has been performed, the trend terms in the maximal fixed model will need to be reparameterized and random trend terms incorporated, as in a) of (ii). Then new random terms will need testing.

In general, the testing of random terms is accomplished using REML Ratio Tests (REMLRTs) when the two models being compared are nested, otherwise Akaike (AIC) and Bayesian (BIC) Information Criteria for the models are compared to determine which has the smaller value. When the REMLRT is a test of whether a component constrained to be nonnegative is zero, the null distribution is a mixture of χ^2s (Self and Liang, 1987). To test fixed effects, conditional Wald F tests are used (Butler et al., 2007).

5. Application of the method to the examples from Piepho et al. (2004)

Example 1: RCBD with one longitudinal factor

The mixed model formulated in Section 4 for this experiment is:

$$\text{GR:}\quad (\mathit{Block}\wedge\mathit{Plot})/\,\text{uc(Lay)};\qquad \text{GF:}\quad \text{(Block+A)}*\text{td(Lay)}.$$

The main difference between it and the model proposed in P4 is that a) the random model explicitly includes the subject term Block∧Plot, and b) it explicitly specifies trends across, and unequal correlation between, the levels of Lay. For the proposed random model, we have that the variance of an observation is $\sigma^2_{\text{BP}}+\sigma^2_{\text{BPL}}$, where σ^2_{BP} corresponds to Block ∧Plot and σ^2_{BPL} to *Block* ∧*Plot* ∧Lay. That is, it is the sum of a component arising from plot variability and another from individual observations. The covariance between two observations on the same plot from layers ℓ and ℓ' is $\sigma^2_{\text{BP}} + \sigma^2_{\text{BPL}}\rho_{\ell\ell'}$, where $\rho_{\ell\ell'}$ is the correlation between two such observations. It is the sum of a) a covariance that is equal between all observations on a plot and reflects the similarity of observations on the same plot, and b) covariance between observations from different layers, on the same plot, that may well differ between the layers. One such model for unequal covariance between layers is that the correlations follow an autoregressive process of order one, referred to by Wolfinger (1993) as an "AR1 plus Common Covariance" structure.

Example 2: RCBD with two longitudinal factors

This example is very similar to the previous one and the result of applying the procedure is:

$$\text{GR:}\quad (\mathit{Block}\wedge\mathit{Plot})/\,\text{uc(Lay}\wedge\text{Date)};\qquad \text{GF:}\quad \text{(Block+A)}*\text{td(Lay)}*\text{td(Date)}.$$

Example 3: Row-column design with one longitudinal factor

Again, this example is a minor variation on the first and the models for it are:

$$\text{GR:}\quad (\mathit{Row}\wedge\mathit{Col})/\,\text{uc(Year)};\qquad \text{GF:}\quad \text{(Row+Col+A)}*\text{td(Year)}.$$

Example 4: RCBD with subsampling and one longitudinal factor

The application of the method for this example is as follows:

IR: ((Block/Plot)∗Year)/Plant;
IF: A.

HR: ⇒ ((Block ∧Plot)/Year)/Plant;
F: A∗Year ⇒ (Block+A)∗Year.

GR: ((Block ∧Plot)/ gf(Year))/Plant ⇒ (*Block* ∧*Plot*)/ uc(Year)/Plant;
GF: (Block+A)∗ td(Year).

Here *Block* ∧*Plot* ∧Year is the only longitudinal error term. Block ∧Plot ∧Year ∧Plant is not, as indicated by Plant being to the right of Year, which it must be as Plant is nested within Year.

Example 5: RCBD with subsampling and two longitudinal factors

This example differs from previous ones in that it involves a randomized, longitudinal factor. The longitudinal factors are Date and Lay, with Date being randomized. Both have subject term Block ∧Plot and Lay also has subject term Block ∧Plot ∧Sample. Application of the method yields:

IR: (Block/Plot/Sample)∗Lay;
IF: A∗Date.

HR: ⇒ (Block ∧Plot)/Lay+(Block ∧Plot ∧Sample)/Lay;
F: A∗Date∗Lay ⇒ (Block+A∗Date)∗Lay.

GR: (Block ∧Plot)/ gf(Date∗Lay)+(Block ∧Plot ∧Sample)/ gf(Lay)
⇒ (*Block*∧*Plot*)/ uc(Date ∧Lay)+(*Block*∧*Plot*∧*Sample*)/ uc(Lay);
GF: (Block+A∗ td(Date))∗ td(Lay).

This model differs from the model in P4 in that it does not include Block ∧Date and Block ∧Date ∧Lay because these are intertier (block-treatment) interactions (Brien and Bailey, 2006, Section 7.1) that are only included in Step 2(a) when there are specific reasons to do so. The interaction between the two unrandomized factors Block and Lay is included a priori, as in P4. Also, two unit terms are included, *Block*∧*Plot* ∧Date ∧Lay and *Block*∧*Plot*∧*Sample* ∧Lay, but the method herein does not include Date in the latter term that contains Sample. One might suppose that only one longitudinal error term should be included to avoid overparameterization. However, it is likely to be possible to estimate their variance parameters, provided there are sufficient data, because the second is involved in covariances on the same day whereas the first is not. On the other hand, there will be models that are overparameterized, such as the compound symmetry model involving these terms, and so to fit them would require the removal of one of them.

Example 6: RCBD with nested longitudinal factor

Here the fixed treatment factor F has three levels of cutting frequencies which are randomized to three different plots in an RCBD. Within each level of F there is a variable number of cuts, the timing of which differs between the levels of F so that the longitudinal factor Cut is nested within F. Not clear in P4 is what actually was cut. Is it the same area, whole or part of the plot, or a different area that was cut each time? If it is the same area, Cut is an unrandomized factor. If not, a number of areas should be designated prior to starting the cutting and Cut randomized to them so that Cut will be a randomized factor. For the case where the same area of the plot is cut, we obtain:

$$\begin{array}{ll} \text{IR:} & \text{Block/Plot/Cut;} \\ \text{IF:} & \text{F/Cut.} \\ & \\ \text{HR:} & (\text{Block}\wedge\text{Plot})/\text{Cut;} \\ \text{F:} & \text{Block} + \text{F/Cut.} \\ & \\ \text{GR:} & (\text{Block}\wedge\text{Plot})/\,\text{gf}(\text{Cut}) \Rightarrow (\textit{Block}\wedge\textit{Plot})/\,\text{uc}(\text{Cut}); \\ \text{GF:} & \text{Block} + \text{F/Cut.} \end{array}$$

This model is the same as in P4, except that Block ∧Cut is not included in the GF model because Cut is always nested. The implied correlation structure between observations on the same plot will depend on which level of F was applied to the Plot. The correlation between levels of Cut for level 2 of F will be the same for all plots that received level 2 of F, but different from the correlation between Cut levels for any other levels of F. Also, it may well be that Cut levels from different levels of F are independent.

Example 7: Split-plot with one longitudinal factor

As for Example 6, the question of whether the same or different areas are cut each time arises. Whichever, as in P4, the longitudinal factor Cut has the two subject terms in *Block ∧Main/Sub*. The application of the method when different cuts are made to the same area of a plot yields:

$$\begin{array}{ll} \text{IR:} & (\text{Block/Main/Sub})*\text{Cut;} \\ \text{IF:} & \text{A}*\text{B.} \\ & \\ \text{HR:} & \qquad\qquad \Rightarrow \text{Block}\wedge\text{Main/Cut} + \text{Block}\wedge\text{Main}\wedge\text{Sub/Cut;} \\ \text{F:} & \text{A}*\text{B}*\text{Cut} \Rightarrow (\text{Block}+\text{A}*\text{B})*\text{Cut.} \\ & \\ \text{GR:} & \text{Block}\wedge\text{Main}/\,\text{gf}(\text{Cut}) + \text{Block}\wedge\text{Main}\wedge\text{Sub}/\,\text{gf}(\text{Cut}) \\ & \quad \Rightarrow \textit{Block}\wedge\textit{Main}/\,\text{uc}(\text{Cut}) + \textit{Block}\wedge\textit{Main}\wedge\textit{Sub}/\,\text{uc}(\text{Cut}); \\ \text{GF:} & (\text{Block}+\text{A}*\text{B})*\,\text{td}(\text{Cut}). \end{array}$$

Example 8: RCBD with random randomized factor and one longitudinal factor

The following shows the use of the method assuming that cuts are to the same area of a plot.

$$\begin{array}{ll} \text{IR:} & (\text{Block/Plot})*\text{Cut;} \\ \text{IF:} & \text{G.} \\ & \\ \text{HR:} & \qquad\quad \Rightarrow \text{Block}\wedge\text{Plot/Cut} + \text{G}\wedge\text{Cut;} \\ \text{F:} & \text{G}*\text{Cut} \Rightarrow \text{Block}*\text{Cut.} \\ & \\ \text{GR:} & \text{Block}\wedge\text{Plot}/\,\text{gf}(\text{Cut}) + \text{G}/\,\text{gf}(\text{Cut}) \\ & \quad \Rightarrow \textit{Block}\wedge\textit{Plot}/\,\text{uc}(\text{Cut}) + \textit{G}/\,\text{uc}(\text{Cut}); \\ \text{GF:} & \text{Block}*\,\text{td}(\text{Cut}). \end{array}$$

In applying Rule 3(a), subject terms have not been restricted to unrandomized factors and this leads to models with unequal correlation associated with G, the alternative model suggested in P4.

Example 9: RCBD at different locations with a random longitudinal factor

In P4, Year is assumed to be random for this experiment with a perennial crop. We believe that this should be done only if it is realistic to do so; it would not be if there was a systematic trend across Year. Random Year is more tenable for an annual crop as new plants are injected into the experiment each year and a systematic trend is less likely. Support for the validity of the assumption of a random longitudinal factor would be provided by a compound symmetry model for the correlation, perhaps with unequal variances, being found suitable. Also, Loc is regarded as fixed, but would more appropriately be random if the differences between treatments vary randomly from one location to another.

The following summarizes the application of the method for this example.

IR: (Loc/Block/Plot)∗Year;
IF: A.

HR: ⇒ (Loc/Block/Plot)∗Year+A ∧(Loc∗Year);
F: A∗Loc∗Year ⇒ A.

GR: (Loc/Block/Plot)∗ gf(Year)+A ∧(Loc∗ gf(Year))
⇒ (*Loc*/*Block*/*Plot*)∗ uc(Year) +A ∧(Loc∗Year);
GF: A.

In this case, subject terms have been restricted to those that involve only unrandomized factors. As usual, our model differs from that in P4 in that the subject terms are included explicitly. As suggested in P4, other models for Years might be entertained. For example, one with fixed trends over Years, that are allowed to vary with A and Loc, as well as random deviations from these trends.

Example 10: RCBD in a crop rotation experiment

Using the terminology of Yates (1954), the experiment involved a single sequence of three crops that was applied three times in three cycles (rotations); a different phase of the sequence was applied to each field so that a field corresponds to a series. In addition we divide the nine years into three triennia. This is appropriate if it is deemed that the years in a triennium are likely to be relatively homogeneous. Although this experiment does not fit our definition of a longitudinal experiment, as the different crops were applied to the field in different years, we derive an analysis for it.

To separate consideration of the natural variation of plots and years from that associated with the imposition of the crop rotation and treatments, we investigate just the unrandomized factors and their relationships. That is, what we would have if no crops or systems had been applied. The unrandomized object is a plot in a year and the unrandomized factors are Field, Block, Plot, Triennium and Year. P4 stated that Year must be nested in Triennium. However, this is not necessarily the case as it depends on whether the first years of the triennia are anticipated to exhibit a degree of similarity; likewise, for the second and third years. However, given that Years is to be considered random, as in P4, nesting of Years within Triennium is appropriate. This means that Triennium on its own is not a longitudinal factor as the years differ between triennia. In addition, Triennium is assumed to be random, which differs from P4 in that in P4 it is assumed that CYC (equivalent to Triennium) is fixed. It seems that Triennium assumed random is consistent with Year assumed random and, as in Example 9, this would be supported if it is established that the appropriate model for the correlation is compound symmetry. If some systematic trend in Triennium is expected, fixed Triennium would be appropriate.

Returning to the assignment of crop rotations and systems, the rotations part involves three Cycles, each of which consists of the same levels combinations of the three factors Phases, Planting and Crop. However, while all three factors are required to specify the Cycle, only one-third of their levels combinations occur, those that do being the same as for a type of 3×3 Latin square. Now the Phase ∧Planting ∧Crop levels are systematically assigned to the Field ∧Year levels within each Triennium, although they could be randomized to Field across Year and Triennium. Also, System is randomized to a Plot within a Block.

In determining the mixed model one needs to take into account that, because a fraction of Phase, Planting and Crop is observed, only a subset of their factorial effects are estimable. However, Crop in combination with Phase is the same as Crop with Planting and only one of the pair is required. Planting allows for systematic differences between years of the rotation and Phase for systematic differences between the phases of a Cycle. Crop∗Planting has the advantage that Planting corresponds to Year; it allows for overall differences between performance of the crops in the three years of a cycle and would be a fixed effect. However, in P4 it assumed implicitly that Planting and Phase have no effect and can be omitted from the analysis. To further understand this experiment it is useful to formulate a decomposition table

Table 2. Decomposition table for crop rotation experiment

plot-year tier		treatments tier	
Source	df	Source	df
Triennium	2		
Year [Triennium]	6	Planting	2
		Planting#Triennium	4
Field	2	Crop#Planting	2
Field#Triennium	4	Crop#Planting#Triennium	4
Field # Year [Triennium]	12	Crop	2
		Crop#Planting	2
		Crop#Triennium	4
		Crop#Planting#Triennium	4
Block [Field]	9		
Block # Year [Field]	18		
Block # Year [Field ∧ Triennium]	54		
Plot [Field ∧ Block]	12	System	1
		System#Crop#Planting	2
		Residual	9
Plot # Triennium [Field ∧ Block]	24	System#Triennium	2
		System#Crop#Planting#Triennium	4
		Residual	18
Plot # Year [Field ∧ Block ∧ Triennium]	72	System#Crop	2
		System#Planting	2
		System#Crop#Planting	2
		System#Crop#Triennium	4
		System#Planting#Triennium	4
		System#Crop#Planting#Triennium	4
		Residual	54

for it and Table 2 is the one based on the inclusion of Crop*Planting in the model. Note that A#B[C ∧ D] stands for the interaction of A and B within the observed combinations of C and D. There is at least one source for each term in the mixed model and the source is obtained from the term as follows: factors in the term that nest any of the factors in the term, and only those factors, must be in the square braces separated by " ∧" and the rest to the left of the square braces separated by "#".

It is clear from Table 2 that separate estimates are not possible for the variance terms a) Field ∧Year ∧ Triennium and Crop ∧Planting ∧Triennium as in P4, b) Field ∧Triennium and Crop ∧Planting ∧ Triennium, c) Year ∧Triennium and Planting ∧Triennium. The terms involving Planting will need to be omitted from the mixed model to produce a model that can be fitted. Also, only two degrees of freedom of Crop#Planting can be estimated as the other two degrees of freedom are inextricably confounded with Field. This would be rectified if Phases were replicated with the incorporation of more Fields.

The following shows the derivation of the mixed model for this experiment:

IR: (Field /Block/Plot)∗(Triennium/Year);
IF: System∗Crop∗Planting∗Triennium.

HR: Field ∧Block ∧Plot∗(Triennium/Year) +Triennium/(Year*(Field/Block)+System∗Crop∗Planting);
F: Field/Block+System∗Crop∗Planting.

GR: Field ∧Block ∧Plot∗(Triennium/Year)
+Triennium/(Year*(Field/Block)+System∗Crop+(System+Crop)∧Planting);
GF: Field/Block+System∗Crop∗ td(Planting).

Another approach would be to assume, if appropriate, no interaction between either or both of Planting and Triennium with Crop and System.

Example 11: RCBD with subplots in strips

To account for all the sources of variability in this experiment, it is necessary to include the factor Row that represents areas that received different irrigations. The factor Irrig is included in the IF model even though they were not randomized to Row because it was assigned, albeit it systematically. Hence the model produced at the end of Stage 1 is not a randomization model. Also, this study does not involve repeated measurements. The factor Irrig corresponds to treatments applied to Row and the Dir to two subareas into which the experimental area is divided:

IR: (Block/Plot)∗Row∗Dir;
IF: Cult∗Irrig.

HR: ⇒ Row/(Block∗Dir)+Row∗((Block ∧Plot)/Dir);
F: Cult∗Irrig∗Dir ⇒ (Block+Cult∗Irrig)∗Dir.

GR: Row/(Block∗Dir)+Row∗((Block ∧Plot)/Dir);
GF: (Block+Cult∗ td(Irrig))∗Dir.

This model is the same as in P4 in so far as Blocks is fixed and it includes the terms Cult ∧Irrig ∧Dir and Block ∧Dir ∧Row. Models other than compound symmetry models may well be appropriate in analyzing this study, but it seems that a more satisfactory basis for these is that they reflect spatial variation over the experimental area rather than serial correlation among Irrig-Dir levels.

One aspect of including the factor Row is that it becomes clear that, because of improper replication (Section 8), overall differences between Irrig levels are inextricably confounded with overall Row differences. Attention has not been drawn to this in previous discussions of the analysis of this experiment.

6. Extension to multitiered experiments

The proposed method extends readily to multitiered experiments (Brien, 1983; Brien and Bailey, 2006), the only change being in determining the intratier models in the first stage. While intratier model formulas are derived as described in Section 4, more than two formulas will be produced for multitiered experiments. These need to be combined to produce the IR and IF models, adding together multiple formulas within a model. The intratier model formula for the observational units will provide the foundation for the IR model formula and formulas for objects that were only randomized will do the same for the IF formula. Other formulas will be incorporated into either model formula as appropriate.

7. The mixed-model analysis of the three-phase example

The proposed method is now applied to the example described in Section 2.

IR and IF models: The observational unit in this experiment is a position and there are four sets of objects: 288 positions, 288 samples, 48 batches, and 6 times. Hence there are four tiers with the factors from each indexing the associated objects; the positions tier is {Runs, Positions}, the samples tier is {Cookings, Samples}, the batches tier is {Kinds, Ages, Lots, Batches} and the times tier is {Times}. Times is a longitudinal factor. The intratier model formulas are Runs/Positions, Cookings/Samples, (Kinds∗Ages)/Lots/Batches, and Times. The first formula corresponds to the observational units and so is included in the IR model, along with the second; the last two formulas are derived from factors that are only ever randomized and so comprise the IF model:

IR: Runs/Positions + Cookings/Samples;

IF: (Kinds∗Ages)/Lots/Batches + Times.

Table 3. Decomposition table for the Eucalypt pulp experiment

positions tier		samples tier		batches-times tiers	
Source	df	Source	df	Source	df
Runs	47	Cookings	47	Kinds	1
				Ages	1
				Kinds#Ages	1
				Lots [Kinds ∧ Ages]	8
				Batches [Kinds ∧ Ages ∧ Lots]	36
Positions [Runs]	240	Samples [Cookings]	240	Times	5
				Times # Kinds	5
				Times # Ages	5
				Times # Kinds # Ages	5
				Times # Lots [Kinds ∧ Ages]	40
				Times # Batches [Kinds ∧ Ages ∧ Lots]	180

HR and F model formulas: Kinds, Ages, and Times will be taken to be fixed factors and all the other factors to be random. The interactions between Times and both Kinds and Ages are of interest. Hence, the HR and F model formulas are:

HR: Runs/Positions + Cookings/Samples + (Kinds ∧Ages ∧Lots)/Batches;

F: Kinds∗Ages∗Times.

To further understand the experiment's analysis a decomposition table showing its confounding is given in Table 3. It was produced using the GenStat procedure `AMTIER` (Brien and Payne, 2006). A table based on just the positions and batches-times tiers could be produced in R or S-PLUS.

GR and GF model formulas: The subject terms for Times are Runs, Cookings, Kinds ∧Ages ∧Lots and Kinds ∧Ages ∧Lots ∧Batches so longitudinal terms involving them and Times will be included in the GR model. Of these, only Runs was never randomized. Unequal correlation between observations with different levels of Times is expected. An initial general random model is

GR: Runs/Positions + *Runs* ∧ uc(Times) + Cookings/Samples + *Cookings* ∧ uc(Times) + (*Kinds* ∧*Ages* ∧*Lots*)/ uc(Times) + (*Kinds* ∧*Ages* ∧*Lots* ∧*Batches*)/ uc(Times).

The random model contains five unit terms and, as is clear from Table 3, several inextricably confounded terms should be removed. The fixed (F) model is modified to allow for trend across the Times. This leads to the following general models with two units terms; let Runs ∧Positions correspond to **R**:

GR: Runs ∧Positions + (*Kinds* ∧*Ages* ∧*Lots*)/ uc(Times)
+ (*Kinds* ∧*Ages* ∧*Lots* ∧*Batches*)/ uc(Times);

GF: Kinds∗Ages∗ td(Times).

The full GR model specifies that the variance of an observation is given by $(\sigma^2_{\mathrm{RP}} + \sigma^2_{\mathrm{CS}}) + (\sigma^2_{\mathrm{RT}} + \sigma^2_{\mathrm{CT}} + \sigma^2_{\mathrm{KALBT}}) + (\sigma^2_{\mathrm{R}} + \sigma^2_{\mathrm{C}} + \sigma^2_{\mathrm{KALB}}) + \sigma^2_{\mathrm{KALT}} + \sigma^2_{\mathrm{KAL}}$ where each σ^2 denotes a particular source of variability arising from the particular combinations of the factors whose initial letters are given in the subscript. For example, $\sigma^2_{\mathrm{KALBT}}$ is the component for variability, between Times of Batches from a particular Lot from one of the Kind-Age combinations. Parenthesized components are not estimable separately. The covariance between two measurements at different times in the same run is given by $(\sigma^2_{\mathrm{RT}} + \sigma^2_{\mathrm{CT}} + \sigma^2_{\mathrm{KALBT}})\rho_{tt'} + \sigma^2_{\mathrm{KALT}}\rho'_{tt'} + (\sigma^2_{\mathrm{R}} + \sigma^2_{\mathrm{C}} + \sigma^2_{\mathrm{KALB}}) + \sigma^2_{\mathrm{KAL}}$ where $\rho_{tt'}$ is the correlation between observations in the same run at times t and t' and $\rho'_{tt'}$ is the correlation between observations from the same lot at times t and t'. The covariance between observations from different runs is $\sigma^2_{\mathrm{KAL}} + \sigma^2_{\mathrm{KALT}}\rho'_{tt'}$ if different Batches from the same Lot are involved, and zero otherwise. In this model $(\sigma^2_{\mathrm{RP}} + \sigma^2_{\mathrm{CS}})$ is a component for measurement error and allows for variability arising from individual positions in the measurement phase and from sampling in the pulp phase. It is the "nugget" variance of geostatistics. A model like this, but with autoregressive correlation of order one, Wolfinger (1993) referred to as "AR1 plus Diagonal".

It may be thought that including the factors Cookings and Samples, and deriving terms from them, is superfluous as these factors are equivalent to Runs and Positions, respectively. Also, Runs appears to be redundant as its levels correspond to the combinations of Kinds, Ages, Lots and Batches. However, their retention draws attention to them all as potential sources of variation and to the need to randomize the factors from one phase to the other.

Table 4. REML ratio tests for random models for the Eucalypt pulp experiment

Model	AIC	BIC	Deviance	REMLRT	df	p
Baseline	997.1	1011.4	988.1			
Random trend modeling						
Reparameterized	1073.3	1109.6	1053.3			
− KALB† spl(T)	1071.6	1104.3	1053.6	0.32	1	0.287^A
− KALB lin(T) + KALB component	1071.2	1100.3	1055.2	1.60	1	0.103^A
− KAL spl(T)	1078.9	1104.4	1064.9	9.74	1	< 0.001
+ KAL spl(T) and intercept-slope covariance	1070.7	1103.4	1052.7	2.52	1	0.112
Covariance modeling for Times	1071.2	1100.3	1055.2			
− KALT	1069.4	1094.8	1055.4	0.18	1	0.337^A
+ unequal T variances‡	1045.4	1089.0	1021.4	33.97	5	$< 0.001^A$
+ AR1 with unequal variances	1028.7	1076.0	1002.7	18.70	1	$< 0.001^A$
− KALB component	1028.0	1072.5	1004.9	2.17	1	0.070
+ AR2 with unequal variances	992.8	1040.0	966.8	38.11	1	$< 0.001^A$
+ AR3 with unequal variances	994.7	1045.6	966.7	0.06	1	0.811
+ unstructured T covariance	962.5	1057.0	910.5	56.26	13	$< 0.001^A$
+ same T variances for A with unequal AT correlations	946.3	1095.3	864.3	46.23	15	$< 0.001^A$
+ T covariances differ for A	953.8	1124.7	859.8	4.46	6	0.614
Check random trend terms						
− KAL lin(T) − KAL spl(T) + KAL component	943.6	1085.3	865.6	1.28	2	0.261^A
− KAL component	950.6	1088.7	874.6	9.02	1	0.001
− A spline	959.4	1097.6	883.4	17.87	1	< 0.001

†Initial letters of some of the factors (K)inds, (A)ges (L)ots, (B)atches and (T)imes.

‡Variances are heterogeneous between Times.

^{A}Change accepted

The different random models were fitted to the data using ASReml-R (Butler et al., 2007). The baseline model was the GR and GF model, but without unequal correlation, the Runs ∧Positions term and trends. Here the trend over time is to be examined and cubic smoothing splines (Verbyla et al., 1999) are to be used, these being appropriate as no functional form for the response had been hypothesized and differences in the trend are of interest. There is evidence in Figure 2 of random variation in these trends and so random trend modeling is performed. The next model to be fitted is the reparameterized mixed model:

$$\text{Kinds}*\text{Ages}*\text{lin(xTime)} \mid \text{spl(xTime)}/(\text{Kinds}*\text{Ages}) + \text{dev(xTime)}/(\text{Kinds}*\text{Ages}) +$$
$$\mathit{Kinds} \wedge \mathit{Ages} \wedge \mathit{Lots} \wedge(\text{lin(xTime)} + \text{spl(xTime)} + \text{Times}) +$$
$$\mathit{Kinds} \wedge \mathit{Ages} \wedge \mathit{Lots} \wedge \mathit{Batches} \wedge(\text{lin(xTime)} + \text{spl(xTime)} + \text{Times})$$

where xTime indicates that the actual values of the levels of Times were used and those terms involving xTime in the random model have variance components, constrained to be positive, fitted for each. There are two types of trend terms in the random model: a) overall trend terms that involve only fixed factors and spl or dev; and b) randomly varying trend terms that involve subject terms and xTime. In this case the only overall trend terms that were not bounded are spl(xTime)/Ages so the rest were dropped. Then REMLRTs were performed to determine the significant randomly varying trend terms and the results are displayed in Table 4. KALB ∧ spl(xTime) was dropped, the slope component of KALB ∧ lin(xTime) replaced with a KALB variance component and the KAL ∧ spl(xTime) and Ages ∧ spl(xTime) terms were retained. Other terms involving xTime were not tested, but covariance between the intercept and slope coefficients for KAL was added and found to be nonsignificant. Attention now turned to covariance modeling for the two longitudinal error terms, although the nonsignificant KALT term is dropped first. A series of models with unequal Time variances and autoregressive correlations between the Times, as well as unstructured covariances between Times, were examined. It was concluded that the best fit to the data was obtained with unstructured covariances between Times, provided that the correlations, but not the variances, differed between the two Ages. A measurement error term was not fitted as it is aliased with the set of unequal variances in the fitted model. The random trend terms were checked and it was found that KAL ∧ spl(xTime) term was now not significant and that the slope component of KAL ∧ lin(xTime) was bounded and can be replaced by a KAL component. Approximate conditional Wald F statistics (Butler et al., 2007) for the fixed terms are shown in Table 5. The fitted mixed model is:

$$\begin{aligned}&\text{Ages} + \text{Kinds}{*}\,\text{lin(xTime)} \mid \text{spl(xTime)/Ages} + \text{Kinds}\wedge\text{Ages}\wedge\text{Lots}\\&+ (\textit{Runs} + \textit{Cookings} + \textit{Kinds}\wedge\text{at}(\textit{Ages})\wedge\textit{Lots}\wedge\textit{Batches})\wedge\text{ush(Times)},\end{aligned}$$

where the "at" operator indicates that the parameters differ for the Ages. From this model it is concluded that the trend over time is nonlinear with different intercepts and curvature for the Ages and different slopes for Kinds. The predicted curves for the different combinations of Age and Kind are shown in Figure 4.

Table 5. Conditional Wald F tests for the fixed effects for the Eucalypt pulp experiment

Fixed term	df numerator	df denominator	Wald F	p
(Intercept)	1	39.2	1076.0	< 0.001
time	1	23.2	19690.0	< 0.001
Kinds	1	9.1	1.2	0.310
Ages	1	38.5	25.4	< 0.001
time:Kinds	1	21.0	4.7	0.004
time:Ages	1	22.7	0.6	0.452
Kinds:Ages	1	9.0	0.3	0.625
time:Kinds:Ages	1	21.8	0.2	0.685

The method in P4 is not strictly applicable in this case because the longitudinal factor Times is randomized. However, suppose we ignore this and that the treatment model is (Kinds∗Ages)/Lots/Batches, the nonrepeated block model is Runs and the repeated factor model is Times. The method in P4 would lead to the mixed model:

$$\text{Kinds}{*}\text{Ages}{*}\text{Times} \mid \text{Kinds}\wedge\text{Ages}\wedge\text{Lots/(Times+Batches)} + \textit{Runs}\wedge\text{Times}.$$

Noting that Runs and Kinds ∧Ages ∧Lots ∧Batches are equivalent, the main difference between this model and that derived above is that unequal correlation models are not considered for the randomized subject term Kinds ∧Ages ∧Lots here. The problem with this is that covariances between Times that differed for Ages would not be considered.

8. Discussion

The differences between the method presented here and P4 are as follows:

(a) The method presented here consists simply of a single sequence of steps in which a mixed model is successively modified and, for it, deciding whether terms are fixed or random is done for all terms

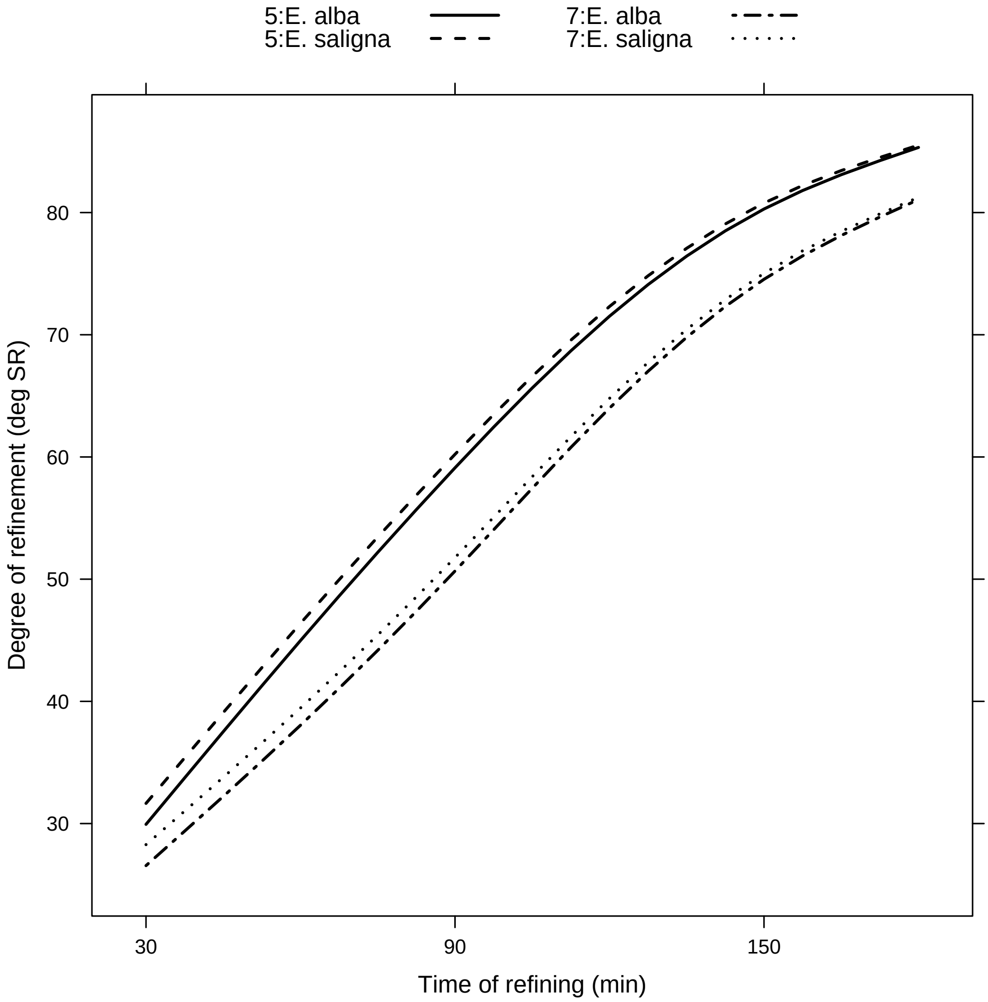

Figure 4. Predicted degree of refinement for each kind and age in the Eucalypt pulp experiment

together. P4 started with several models and combined them in a complicated manner; terms were designated as fixed or random separately for different models.

(b) The extension from Brien and Bailey (2006) to the present method is transparent so that it is obvious that the extension provides a general method of formulating mixed models for both experiments that are longitudinal and those that are not. Applying the method of P4 to nonlongitudinal experiments produces only ANOVA-type mixed models.

(c) Here a mixed model equivalent to the randomization model is produced during the formulation, in Stage 1, but not for P4.

(d) Unlike in P4, separate subject terms are always included, at least initially, for reasons explained in Section 4. In this method, unequal correlation models are considered for randomized subject terms, but are not in P4.

(e) Here, but not in P4, trend and unequal correlations operators are included explicitly in model formulas and subject specific effects are initially stipulated in the model.

(f) Longitudinal factors are allowed to be randomized here, unlike in P4.

(g) In the method presented here, multitiered experiments are covered, whereas P4 is restricted to the subset of multiphase experiments.

8.1 Why randomization-based models?

Stage 1 of our method is important because it ensures that all the terms, taken into account in the randomization, are included in the analysis and that the incorporation of any extra terms not taken into account is intentional. In this regard we strongly recommend against ever omitting unrandomized factors from the analysis by replacing them with randomized factors, as in Rule 5 in P3 and as is done by Littel et al. (2006, Section 4.2). Doing so leads to a misidentification of sources of variation, including the possible omission of the experimental units as sources. For example, for the RCBD, the mixed model equivalent to the randomization model is Treatments | Blocks + Blocks $\wedge$Plots. Rule 5 modifies this to Treatments | Blocks+ Blocks $\wedge$Treatments. Of course, the latter is economical because the factor Plots is no longer required. However, whereas the original model included the term Blocks $\wedge$Plots, whose levels are the experimental units, the latter does not. Certainly, the combinations of Blocks $\wedge$Treatments are not the experimental units in that Treatments are not applied to the combinations of Blocks $\wedge$Treatments. This gives rise to the difficulty, addressed by Samuels, Casella, and McCabe (1991), of distinguishing between what they termed "independent contributions", or sources of error variation, and interactions between factors. In the case of the RCBD, the modified model gives the false impression that the second source of variation is the interaction of Blocks and Treatments, whereas under the assumption of Block and Treatment additivity, it arises from the differences between Plots within Blocks as is obvious from the randomization model. Similarly, the mixed model equivalent to the randomization model for the split plot, with factor A randomized to main plots with an RCBD and B randomized to subplots, is A $*$ B | Blocks + Blocks $\wedge$Plots + Blocks $\wedge$Plots $\wedge$Subplots. This makes clear that the main plot error is also the natural or error variation arising from differences between Plots within Blocks and avoids the confusion that Samuels et al. (1991) discussed. Rule 5 also has the unfortunate effect of promoting the presentation of data in treatment order so that the randomized order is frequently lost. Consequently, some forms of diagnostic checking and later reanalysis are impossible.

However, the models are randomization-based only to the extent that the randomization is used as a basis for determining the terms in the model, and not whether terms are fixed or random nor as the basis for statistical inference. The reasons for this are that there are many situations in which a) inferences cannot be randomization based or b) a mixed model that differs from the intratier mixed model, in ways allowed for in Stages 2 and 3, better fits the data. Thus, there is no randomization justification for inferences about intertier interactions so that a mixed model must be used when they are of interest. A particular case are interactions between unrandomized longitudinal factors and randomized factors. More controversial is when the longitudinal factors are randomized, as in the example in Section 7. Indeed, P4 suggested that in these cases one does not need to employ serial correlation models as the randomization justifies the assumptions of equal correlation and homogeneous variance. It is our contention that often the longitudinal factors induce unequal correlation and heterogeneous variance amongst the units to which they have been randomized and this needs to be accounted for in the analysis by using mixed models that reflect this.

Another variation from the intratier model is the designation of unrandomized factors as fixed, like Lay and Block in Example 1 and Loc in Example 9, rather than random as implied by the randomization models. Clearly, this is necessary whenever trends over longitudinal factors, like Lay, are anticipated. Similarly, if blocks are contiguous with a smooth trend, then a model with a trend over Block and random deviations will fit better than one with random Block. In this case a general model for Example 1 might be: (lin(Block)+A)$*$ td(Lay) | spl(Block) + dev(Block) + $(Block \wedge Plot)$/ uc(Lay). However, a model with Block fixed has the virtue of not requiring normally distributed Block effects and so involves less assumptions.

8.2 Confounding and improper replication

A primary consequence of randomization is the confounding between terms. As we have seen in Example 10 and in Section 7, the decomposition table for an experiment, when based on the randomization, displays this confounding. It is particularly useful for detecting situations in which some terms are inextricably confounded so that separate estimates are not possible. This is required when mixed model software is being employed as all but one of a set of inextricably confounded terms needs to be omitted from the model for it to be successfully fitted. That this is necessary is a disadvantage of mixed model software as it means that the model supplied is a "model of convenience" because terms for potential sources of

Table 6. Decomposition table for the improperly replicated continuous grazing trial

animals tier		plots tier		treatments tier	
Source	df	Source	df	Source	df
Classes	$a-1$				
Animals [Classes]	$a(t-1)$	Plots	$t-1$	Treatments	$t-1$
		Residual	$(a-1)(t-1)$		

variation are omitted from the full model for the sake of obtaining a fit. Mixed models of convenience are particularly dangerous when there is improper replication involved.

Brien and Demétrio (1998) described an improperly replicated continuous grazing trial in which t treatments are applied to t plots. In addition, ta animals are divided into a homogeneous weight classes based on initial weight, and the t plots are randomized to the t animals in each weight class. This will result in a animals assigned to each plot. The decomposition table corresponding to the analysis for an animal response variable is given in Table 6. Now, it is clear the Plots and Treatments are inextricably confounded and Plots will need to be omitted from the mixed model for this experiment, which would be: Treatments | Classes/Animals. The problem with this model is that there is no indication that the test for Treatments may be misleading because a significant result may be due to Plot variation and not to Treatment differences. The improper replication may well go unnoticed.

Acknowledgments

The authors would like to thank CIAM, University of South Australia and CAPES, CNPq, FAPESP, Brasil for their assistance in making possible the exchange visits of the authors to work on this article.

REFERENCES

Bailey, R. A. (1981). A unified approach to design of experiments. *Journal of the Royal Statistical Society, Series A 144*, 214–223.

Brien, C. J. (1983). Analysis of variance tables based on experimental structure. *Biometrics 39*, 51–59.

Brien, C. J. and R. A. Bailey (2006). Multiple randomizations (with discussion). *Journal of the Royal Statistical Society, Series B 68*, 571–609.

Brien, C. J. and R. A. Bailey (2009). Decomposition tables for experiments. I. A chain of randomizations. *The Annals of Statistics*. accepted for publication.

Brien, C. J. and C. G. B. Demétrio (1998). Using the randomisation in specifying the ANOVA model and table for properly and improperly replicated grazing trials. *Australian Journal of Experimental Agriculture 38*, 325–334.

Brien, C. J. and R. W. Payne (2006). AMTIER Procedure. In R. W. Payne and P. W. Lane (Eds.), *GenStat Reference Manual Release 9, Part 3. Procedure Library PL 17*, pp. 73–75. Hemel Hempstead: VSN International.

Butler, D., B. R. Cullis, A. R. Gilmour, and B. J. Gogel (2007). *Analysis of Mixed Models for S language environments: ASReml-R reference manual.* Brisbane: DPI Publications.

Fisher, R. A. (1935). Contribution to the discussion of "Complex experiments" by F. Yates. *J. R. Statist. Soc. Suppl. 2*, 229–231. Incomplete extract at URL: <http://hdl.handle.net/2440/15225>, (accessed July 9, 2008).

Hurlbert, S. H. (1984). Pseudoreplication and the design of ecological field experiments. *Ecological Monographs 54*, 187–211.

Insightful Corporation (2005). *S-PLUS 7.0 for Windows.* Seattle, Washington: Insightful Corporation.

Laird, N. M. and J. H. Ware (1982). Random-effects models for longitudinal data. *Biometrics 38*, 963–974.

Littel, R., G. Milliken, W. Stroup, R. Wolfinger, and O. Schabenberger (2006). *SAS for Mixed Models* (2nd ed.). Cary: SAS Press.

Lorenzen, T. and V. L. Anderson (1993). *Design of Experiments: A No-Name Approach.* New York: Dekker.

Nelder, J. A. (1965a). The analysis of randomized experiments with orthogonal block structure. I. Block structure and the null analysis of variance. *Proceedings of the Royal Society, Series A 283*, 147–161.

Nelder, J. A. (1965b). The analysis of randomized experiments with orthogonal block structure. II. Treatment structure and the general analysis of variance. *Proceedings of the Royal Society, Series A 283*, 162–178.

Patterson, H. D. (1997). Analyses of series of variety trials. In R. A. Kempton and P. N. Fox (Eds.), *Statistical Methods for Plant Variety Evaluation*, pp. 139–161. London: Chapman and Hall.

Payne, R. W., S. A. Harding, D. A. Murray, D. M. Soutar, D. B. Baird, A. I. Glaser, I. C. Channing, S. J. Welham, A. R. Gilmour, R. Thompson, and R. Webster (2008). *The Guide to GenStat Release 11, Part 2 Statistics.* Hemel Hempstead: VSN International. URL: <`http://www.genstat.co.uk/resources/documentation/genstat-guide-statistics/`>, (accessed July 9, 2008).

Pereira, R. A. G. (1969). *Estudo Comparativo das Propriedades Físico-Mecânicas da Celulose Sulfato de Madeira de Eucalyptus saligna* Smith*, Eucalyptus alba* Reinw *e Eucalyptus grandis* Hill ex Maiden. Ph. D. thesis, Escola Superior de Agricultura 'Luiz de Queiroz', University of São Paulo, Piracicaba, Brasil.

Piepho, H. P. (2006). Contribution to the discussion of "Multiple randomizations" by C. J. Brien and R. A. Bailey. *Journal of the Royal Statistical Society, Series* B *68*, 605.

Piepho, H. P., A. Büchse, and K. Emrich (2003). A hitchhiker's guide to mixed models for randomized experiments. *Journal of Agronomy and Crop Science 189*, 310–322.

Piepho, H. P., A. Büchse, and C. Richter (2004). A mixed modelling approach for randomized experiments with repeated measures. *Journal of Agronomy and Crop Science 190*, 230–247.

Pinheiro, J. C. and D. Bates (2000). *Mixed Effects Models in S and S-PLUS.* New York: Springer.

R Core Development Team (2008). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. URL: <`http://www.r-project.org`>.

Robinson, G. K. (1991). That BLUP is a good thing: The estimation of random effects. *Statistical Science 6*, 15–51.

Samuels, M. L., G. Casella, and G. P. McCabe (1991). Interpreting blocks and random factors. *Journal of the American Statistical Association 86*, 798–821.

Self, S. G. and K.-Y. Liang (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association 82*, 605–610.

Verbeke, G. and G. Molenberghs (2000). *Linear Mixed Models for Longitudinal Data.* Berlin: Springer-Verlag.

Verbyla, A. P., B. R. Cullis, M. G. Kenward, and S. J. Welham (1999). The analysis of designed experiments and longitudinal data by using smoothing splines (with discussion). *Applied Statistics 48*, 269–311.

White, R. F. (1975). Randomization and the analysis of variance. *Biometrics 31*, 552–572.

Wilk, M. B. and O. Kempthorne (1956). Some aspects of the analysis of factorial experiments in a completely randomized design. *Annals of Mathematical Statistics 27*, 950–985.

Wolfinger, R. D. (1993). Covariance structure selection in general mixed models. *Communications in Statistics, Simulation and Computation 22*, 1079–1106.

Yates, F. (1954). The analysis of experiments containing different crop rotations. *Biometrics 10*, 324–346.

A. Data for the Eucalypt pulp experiment

					Time	30	60	90	120	150	180
Run	Kinds	Ages	Lots	Batches	Position	1	2	3	4	5	6
1	E. saligna	5	1	1		29.0	46.0	51.0	74.5	80.5	85.5
2	E. saligna	5	1	2		36.0	50.0	67.5	73.5	82.5	83.5
3	E. saligna	5	1	3		32.0	48.5	63.5	71.5	79.0	86.0
4	E. saligna	5	1	4		32.5	47.0	61.0	69.0	79.0	84.5
5	E. saligna	5	2	1		32.0	45.0	61.0	70.0	82.0	86.5
6	E. saligna	5	2	2		26.0	40.0	55.0	70.5	78.5	84.0
7	E. saligna	5	2	3		30.0	43.0	60.0	80.0	82.0	87.5
8	E. saligna	5	2	4		35.0	51.0	71.0	79.5	88.0	91.0
9	E. saligna	5	3	1		36.0	52.0	69.5	78.5	86.5	86.5
10	E. saligna	5	3	2		26.5	40.5	56.5	67.5	79.5	83.5
11	E. saligna	5	3	3		36.5	56.0	67.0	81.0	89.0	89.0
12	E. saligna	5	3	4		32.0	47.0	64.0	78.0	87.0	89.0
13	E. saligna	7	1	1		29.0	42.0	46.0	66.5	74.5	82.0
14	E. saligna	7	1	2		25.5	32.0	43.0	55.0	65.5	77.0
15	E. saligna	7	1	3		28.0	39.0	42.0	49.0	61.0	73.5
16	E. saligna	7	1	4		26.0	40.0	48.0	69.5	79.5	83.5
17	E. saligna	7	2	1		26.0	43.0	45.0	71.0	78.0	81.0
18	E. saligna	7	2	2		24.5	47.5	47.0	73.0	71.0	80.0
19	E. saligna	7	2	3		25.0	34.0	46.5	62.0	70.0	77.0
20	E. saligna	7	2	4		29.0	39.0	52.0	67.0	77.0	84.0
21	E. saligna	7	3	1		28.0	37.0	48.0	68.0	75.5	82.0
22	E. saligna	7	3	2		27.5	38.0	54.0	71.0	79.0	83.0
23	E. saligna	7	3	3		29.0	40.5	55.0	71.0	80.0	84.5
24	E. saligna	7	3	4		31.0	42.0	58.0	68.5	80.0	85.5
25	E. alba	7	1	1		23.0	44.0	51.0	61.0	72.5	81.5
26	E. alba	7	1	2		27.0	44.0	60.0	75.0	80.0	84.5
27	E. alba	7	1	3		28.0	43.5	51.0	70.5	80.5	81.5
28	E. alba	7	1	4		30.0	46.0	64.0	72.0	79.5	85.5
29	E. alba	7	2	1		29.0	41.0	53.0	65.0	77.0	85.0
30	E. alba	7	2	2		31.0	44.0	60.0	74.5	82.5	87.5
31	E. alba	7	2	3		32.0	46.0	61.0	73.0	80.5	84.5
32	E. alba	7	2	4		31.0	46.0	60.0	72.5	80.5	87.5
33	E. alba	7	3	1		29.5	40.0	52.0	61.0	72.0	80.0
34	E. alba	7	3	2		33.5	45.5	62.0	73.0	81.0	86.0
35	E. alba	7	3	3		30.0	44.0	61.0	74.5	83.5	88.5
36	E. alba	7	3	4		33.0	49.0	65.5	75.5	86.0	89.0
37	E. alba	5	1	1		23.0	31.0	43.0	52.0	65.5	77.5
38	E. alba	5	1	2		25.0	39.0	47.0	59.0	72.0	80.0
39	E. alba	5	1	3		21.0	30.0	44.0	49.0	65.5	77.5
40	E. alba	5	1	4		22.0	30.0	39.0	47.0	62.0	73.5
41	E. alba	5	2	1		28.0	38.0	54.0	69.0	76.5	82.0
42	E. alba	5	2	2		32.5	39.5	57.0	77.0	84.0	83.0
43	E. alba	5	2	3		26.5	52.0	52.5	77.0	83.0	84.0
44	E. alba	5	2	4		27.0	40.0	58.0	67.0	77.0	83.0
45	E. alba	5	3	1		28.0	40.5	54.0	68.0	78.5	84.5
46	E. alba	5	3	2		28.0	38.0	51.0	64.0	74.0	81.0
47	E. alba	5	3	3		30.0	43.0	62.0	73.5	84.5	86.5
48	E. alba	5	3	4		28.0	40.0	55.0	68.5	76.5	82.5
49	E. grandis	7	1	1		30.0	45.0	59.0	72.5	80.5	83.5
50	E. grandis	7	1	2		30.0	43.0	59.0	72.0	79.5	84.0
51	E. grandis	7	1	3		29.5	42.5	57.0	69.5	77.5	82.5
52	E. grandis	7	1	4		29.5	42.5	57.5	68.5	80.5	82.5
53	E. grandis	7	2	1		35.0	49.0	66.5	77.5	86.0	91.0
54	E. grandis	7	2	2		32.0	48.0	64.0	75.5	85.5	90.0
55	E. grandis	7	2	3		32.0	47.0	64.5	75.5	85.5	89.5
56	E. grandis	7	2	4		32.0	48.0	66.0	78.0	86.0	91.0
57	E. grandis	7	3	1		34.0	51.0	68.5	77.5	85.5	83.5
58	E. grandis	7	3	2		26.0	47.0	54.0	77.5	84.5	83.5
59	E. grandis	7	3	3		25.0	36.0	51.0	66.0	78.5	85.0
60	E. grandis	7	3	4		31.0	49.0	66.5	79.0	87.0	90.0