J. Agronomy & Crop Science **189**, 310—322 (2003)

ISSN 0931-2250

Department of Bioinformatics, Institute of Crop Production and Grassland Research, University Hohenheim, Fruwirthstraße, Stuttgart, Germany

A Hitchhiker's Guide to Mixed Models for Randomized Experiments

H. P. Piepho, A. Büchse and K. Emrich

Authors' address: Prof. Dr H. P. Piepho (corresponding author; e-mail: piepho@uni-hohenheim.de), Dr A. Büchse and Dr K. Emrich, Department of Bioinformatics, Institute of Crop Production and Grassland Research, University Hohenheim, Fruwirthstraße 28, 70599 Stuttgart, Germany

With 1 table and 1 figure

Received February 19, 2003; accepted March 3, 2003

Abstract

Designed experiments conducted by crop scientists often give rise to several random sources of variation. Pertinent examples are split-plot designs, series of experiments and repeated measurements taken on the same field plot. Data arising from such experiments may be conveniently analysed by mixed models. While the mixed model framework is by now very well developed theoretically, and good software is readily available, the technology is still underutilized. The purpose of the present paper is, therefore, to encourage more widespread use of mixed models. We outline basic principles, which help in setting up mixed models appropriate in a given situation, the main task required from users of mixed model software. Several examples are considered to demonstrate key issues. The theoretical underpinnings are briefly sketched in so far as they are practically relevant for making informed use of mixed-model computer packages. Finally, a brief review is given of some recent methodological developments, which are of interest to the plant sciences. A German version of this paper is available from the corresponding author upon request.

Key words: analysis of variance — blocking — error strata — geostatistics — longitudinal data — mixed model — random effect — randomization — repeated measurements — series of experiments

Introduction

In agricultural research, designed experiments are usually analysed based on a linear model. More often than not, the model is of the mixed type, because it includes several effects representing different random sources of variation. A few examples are as follows: (1) A split-plot experiment requires two error terms for main-plots and subplots. (2) Repeated measurements taken at different points in time and/or space are correlated, which may be accounted for by a mixed model with an appropriate variance–covariance structure. (3) The spatial analysis of geostatistical data is readily embedded in a mixed model framework. (4) Experiments replicated at several sites and/or in several years call for a linear model with random environmental effects, providing environments may be considered as a random factor and the objective is to compute treatment means across environments. (5) Recovery of inter-block information in an incomplete block design can be invoked by specifying a random block effect.

Prior to the advent of computers, the analysis of a mixed model was a daunting task. In fact, analysis was only feasible for simple, balanced designs. A full-fledged analysis of more complex data sets, e.g. of an unbalanced series of experiments accommodating heterogeneity of variance at various levels (treatment by environment interaction and error) and spatial correlation at the field level (Smith et al. 2001), was beyond reach. While the mixed model framework has been well developed over the past 20 years (McLean et al. 1991, Searle et al. 1992, Verbeke and Molenberghs 2000) and analyses are now easy to perform with modern statistical software, the use of mixed models in agricultural research does not seem to have reached the level it deserves. This discrepancy is partly due to the fact that presentation of the methodology in many textbooks is rather heavy on the theoretical side, making mixed models seem more difficult than they actually are. Also, the output of mixed model packages is somewhat unfamiliar to those accustomed to analysis-of-variance (ANOVA) tables and least significant differences, although there are far more similarities than dissimilarities.

We are convinced that agricultural scientists can produce valid and useful mixed model analyses with little difficulty if equipped with the appropriate software and a solid understanding of some basic principles. The purpose of the present paper is to describe these principles. In so doing, we hope to help disseminate this powerful and versatile methodology. A major focus will be on the formulation of a mixed model, the main task in a computer-based analysis. We also give a cursory review of some recent developments in mixed model methodology which are relevant to plant scientists. Readers already fluent in mixed models will find little new material. Our intended readership are researchers with a working background in classical ANOVA, who want to get a grip on mixed models.

A mixed model analysis consists of two major steps: (1) Setting up a model. (2) Making statistical inferences (parameter estimation, tests and confidence intervals). Our paper devotes a section to each of these two steps. In 'Rules for Setting up a Mixed Model', we give a number of rules we have found useful in developing mixed models. The rules are focused on ANOVA-type mixed models for qualitative treatment factors (Hocking 1985). The principles of statistical inference are discussed in 'Statistical inference'. Our model notation is described in 'Rules for Setting up a Mixed Model'. Extensions of the ANOVA-type mixed model will be briefly reviewed at the end of the paper (see 'A brief review of some extensions to ANOVA-type mixed models as relevant in agronomy'). The presentation is written with the user of a statistical package in mind. There are many good packages for mixed model analysis (ASREML, BMDP, GENSTAT, SAS, S-PLUS and SPSS). While the principles we outline apply to any of these packages, for illustration we occasionally refer to the SAS system (Littell et al. 1996), a package with which we are familiar and which is quite commonly used among agronomists. We would like to stress, however, that our referring to one particular package does not imply a specific recommendation.

Model Syntax

A convenient model notation

Example 1

A split-plot experiment is performed in which the main-plot factor A (fertilizer) is laid out in randomized complete blocks (R), while the sub-plot factor B (genotype) is completely randomized within main-plots. A linear model for the data can be written as

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + r_k + b_{ik} + e_{ijk} \quad (1)$$

where y_{ijk} is the yield of ith fertilizer and jth genotype in kth block, μ the general mean, α_i the main effect of ith fertilizer, β_j the main effect of jth genotype, $(\alpha\beta)_{ij}$ the fertilizer-by-genotype interaction, r_k the effect of kth block, b_{ik} the error of ith main-plot within kth block and e_{ijk} the sub-plot error (residual).

All effects are fixed, except for the random error terms b_{ik} and e_{ijk}, which are assumed to be normally distributed with zero mean and variances σ_b^2 and σ_e^2, respectively. We consider (1) an ANOVA-type mixed model because of the simple assumption of independence and homogeneity of variance for all random effects. There are numerous extensions to ANOVA-type mixed models, which relax these assumptions. A brief review will be given at the end of the paper.

Assume that in a data file to be submitted to analysis, genotypes are coded by A, fertilizers are coded by B and blocks are coded by R. Then an alternative expression for the above model, which is more akin to the syntax used by statistical packages for linear models, is

$$\mathrm{Y} = \mathrm{A} + \mathrm{B} + \mathrm{A}\cdot\mathrm{B} + \mathrm{R} + \mathrm{R}\cdot\mathrm{A} + \mathrm{R}\cdot\mathrm{A}\cdot\mathrm{B} \quad (2)$$

with the following identities: Y equals y_{ijk}, A equals α_i, B equals β_j, A · B equals $(\alpha\beta)_{ij}$, R · A equals b_{ik} and R · A · B = e_{ijk}. Model (2) does not contain a symbol for the general mean or intercept, μ, based on the premise that (virtually) any linear model contains this effect. Often, the residual term (R · A · B in this case) is not stated either. In this paper we retain and underline the residual for clarity. To further simplify the notation, one may drop the response Y (Wilkinson and Rogers 1973), yielding

$$\mathrm{A} + \mathrm{B} + \mathrm{A}\cdot\mathrm{B} + \mathrm{R} + \mathrm{R}\cdot\mathrm{A} + \underline{\mathrm{R}\cdot\mathrm{A}\cdot\mathrm{B}}. \quad (3)$$

The effects A, B, A · B and R are fixed, while R · A and $\underline{\mathrm{R}\cdot\mathrm{A}\cdot\mathrm{B}}$ are random. This may be indicated by separating fixed and random effects with a colon, listing fixed effects first (Patterson 1997):

$$\mathrm{A} + \mathrm{B} + \mathrm{A}\cdot\mathrm{B} + \mathrm{R} : \mathrm{R}\cdot\mathrm{A} + \underline{\mathrm{R}\cdot\mathrm{A}\cdot\mathrm{B}}. \quad (4)$$

We use the notation in (4) throughout for two reasons: (i) it is simpler than the notation in (1) and yet conveys basically the same information; (ii) it is

essentially the notation used with statistical software. For 'Example 1' the important statements with the MIXED procedure of the SAS system would be

$$\text{model Y} = \text{A B A} * \text{B R};$$

$$\text{random R} * \text{A};$$

All fixed effects are listed in the *model* statement, while all random effects, except the residual, are specified in the *random* statement. The residual need not be stated explicitly, as it will be fitted automatically.

Four model operators

To derive a model appropriate for a given design, we use a short-hand notation proposed by Nelder (1965), Wilkinson and Rogers (1973) and McCullagh and Nelder (1989). This notation involves four operators, which we now describe. The operators provide a convenient way of stating a model in compact form. The notation is intuitively appealing and quickly reveals the main features of the design.

Dot operator (·)
The dot operator is used to define crossed effects. The formation of a crossed effect may be considered as a mathematical operation, just as the addition of terms in a model using the '+' operator. The dot operator has higher priority, i.e. $A \cdot (B + C) = A \cdot C + A \cdot B$. Also, the dot operator is commutative and associative, so that $A \cdot B = B \cdot A$ and $(A \cdot B) \cdot C = A \cdot (B \cdot C)$. When two effects are crossed, which contain the same factor, duplicate terms are deleted. For example, $(R \cdot A) \cdot (R \cdot B) = R \cdot A \cdot B$.

Product-term operator [pt(.)]
If M denotes a model, then pt(M) is the product term (using dots) of all effects in M. For example, $\text{pt}(A + B) = A \cdot B$ and $\text{pt}(A + B \cdot C) = A \cdot B \cdot C$.

Nesting operator (/)
If a factor B is nested within another factor A, the model must contain the terms A and $A \cdot B$. A short-hand notation for this is A/B, where '/' is a nesting operator. Thus, we may write $A/B = A + A \cdot B$. If an additional factor C is nested within B, we may write A/B/C. The nesting operator is associative, i.e. $A/(B/C) = (A/B)/C$. Moreover, we have $A/(B + C) = A/B + A/C$. If A and B are themselves model formulae, A/B is defined as $A/B = A + \text{pt}(A) \cdot B$.

Crossing operator (×)
The model for two crossed factors A and B can be represented by $A \times B = A + B + A \cdot B$. We use '×' in place of the symbol '*' proposed by McCullagh and Nelder (1989) to avoid confusion with the SAS syntax, in which '*' denotes the dot operator. Two important algebraic rules for a set of three factors A, B and C are

$$\begin{aligned} A \times (B + C) &= A + (B + C) + A \cdot (B + C) \\ &= A + B + C + A \cdot B + A \cdot C \end{aligned} \quad (5)$$

and

$$\begin{aligned} (A \times B)/C &= A \times B + \text{pt}(A \times B) \cdot C \\ &= A \times B + A \cdot B \cdot C. \end{aligned} \quad (6)$$

Rules for Setting up a Mixed Model

In this section, we state and exemplify a number of rules, which help in setting up a mixed model. For ease of reference, an overview of the rules, together with the Journal page, is given in Table 1.

Random vs. fixed

Rule 1 (when is a factor random?)
A factor is random when the observed levels can be regarded as randomly sampled from a population (e.g. environments and sampling units). Alternatively, a factor is random if it represents a randomization unit (e.g. plots). Otherwise the factor is usually taken as fixed (e.g. non-randomized blocks and treatments). If comparisons are to be made among the levels of a factor (e.g. treatments), the factor is considered as fixed, regardless of whether or not it is random by

Table 1: Summary of rules for setting up an ANOVA-type mixed model

Rule	Key phrase	Journal page
1	When is a factor random?	312
2	Two types of factor	313
3	Keep treatment and block model separate, at least initially	313
4	Effects of the block model	313
5	Coding of block factors	314
6	Interaction among block and treatment factors	315
7	Multi-phase experiments	315
8	Taking random effects fixed	316

design. If a factor is random, then all effects containing that factor are random.

Example 2
A series of experiments is performed with a set of selected genotypes. Experiments are replicated across a random sample of locations from a target region. The objective of the analysis is to obtain genotype means across locations representative of the target region. The design entails two factors: genotypes and locations. Genotype is a fixed factor, because comparisons among levels of that factor are of interest and because the genotypes were selected, i.e. they do not represent a random sample of some well-defined population of genotypes. By contrast, location is a random factor because locations were randomly sampled and there is no interest in the levels sampled; test locations merely serve as replications. The interaction of genotype and environment is random, because it contains the location factor.

It should be pointed out that quite often, locations are not selected at random, e.g. when existing research facilities are used, which were placed to represent certain agroecological zones. In these cases, it is more appropriate to regard locations as a fixed factor.

Two types of factor

Rule 2 (two types of factor)
We may distinguish treatment and block factors. *Block* factors comprise randomly selected sampling units (plants, soil samples, etc.), randomization units (rows, columns, incomplete blocks, main-plots, sub-plots, etc.) and blocking units not involved in randomization (complete blocks, environments, etc.). Block factors are needed to uniquely identify each *observational unit* (plot, plant, etc.), i.e. block factors are *innate* to the observational units (Brien 1983). A *treatment* factor and its levels are chosen by the experimenter to answer a research question. Levels of a randomized treatment factor are allocated to observational units by a randomization process, i.e. treatments are *not innate* to observational units.

Example 3
A series of randomized complete block experiments with different genotypes is replicated across locations. Genotype is a randomized treatment factor, while location, complete block and plot are block factors. Specifically, a location may be regarded as a super-block made up of several complete blocks. The block factors are needed to identify and are innate to the observational units (plots). By contrast, genotypes are allocated to observational units by a randomization process and so are not innate to them.

Treatment and block model for a randomized experiment

The rules we give in this section are essentially those proposed by Nelder (1965) and Wilkinson and Rogers (1973). They are also related to directives for obtaining the block model as implemented in GENSTAT (Payne and Wilkinson 1977) and GLIM (McCullagh and Nelder 1989). The rules are based on R. A. Fisher's premise of no interaction among treatment and block effects (Nelder 1965, Brien 1983, Bailey 1991).

Rule 3 (keep treatment and block model separate, at least initially)
In modelling a designed experiment, it is useful to keep the treatment model and the block model separate (Nelder 1965). The treatment structure can be modelled solely using treatment factors, while the block model can be expressed exclusively in terms of block factors. The block model fully describes the data when there are no treatment effects and it represents the structure innate to the observational units.

Rule 4 (effects of the block model)
In a designed experiment, each randomization unit (error stratum) receives a separate random effect. An experimental or blocking unit becomes a randomization unit, when levels of a factor or factor combination (treatment or other) are randomly allocated to different units. If two randomizations (error strata) are *crossed*, i.e. a unit of one randomization extends across several units of the other, the cross generates a further experimental unit (e.g. a plot). The model, therefore, also contains a random effect obtained by crossing the two variables representing the crossed randomization units. If more than two randomizations are mutually crossed, the model contains random effects obtained by all possible crosses among the variables representing the crossed randomization units (two-way, three-way, etc.). Each type of sampling unit also receives a separate random effect. Finally, the block model contains fixed effects for blocking factors, which are not involved

in randomization or sampling, e.g. blocks in a randomized complete block design. The random effect corresponding to the observational unit is equivalent to the residual.

Rule 5 (coding of block factors)
Each block factor can be uniquely coded by a separate variable. This coding makes it easy to set up a correct block model. After the block model has been formulated, a variable coding for a block factor can often be replaced by a variable coding for a treatment factor or by a crossed effect of several treatment factors. In this event, the variable coding the block factor may vanish from the final model. It should be stressed, however, that this type of replacement is not needed for obtaining a correct analysis, and one may prefer not to make it, particularly in complex experiments, where it can obscure the main features of the design (Brien 1983).

Example 4
A treatment factor A is tested in randomized complete blocks (R). The block model comprises a fixed effect for complete blocks and a random effect for plots (randomization units) *nested* within complete blocks. Assume that the plots within a block are uniquely coded by the variable PLOT, i.e. each plot in a block has a different level of PLOT. The block model has the form $R : \underline{R \cdot PLOT}$, while the treatment model is A. A short-hand for the *nested* randomization of plots within blocks is R/PLOT. The complete model is $A + R : \underline{R \cdot PLOT}$. A plot within a block is also uniquely identified by the level of A tested on the plot, so we may replace PLOT by A and use the equivalent model $A + R : \underline{R \cdot A}$.

Example 5
A treatment factor A is tested in a Latin square design. There are two *crossed* randomization units, i.e. rows (ROW) and columns (COL). Hence, the block model is $ROW \times COL = ROW + COL + \underline{ROW \cdot COL}$. The crossed effect ROW · COL represents the observational unit (plot), which is generated by the crossing of rows and columns. Adding the treatment model, the full model reads $A : ROW + COL + \underline{ROW \cdot COL}$.

Example 6
A split-plot experiment with two factors A and B is conducted with main-plots arranged in complete blocks (R). Factor A is the main-plot factor, i.e. levels of A are randomly allocated to main-plots within a block. Levels of B are randomly allocated to sub-plots within a main-plot. There are two *nested* randomization units: main-plots (MAIN) within complete blocks and sub-plots (SUB) within main-plots. A short-hand for the block model is R/MAIN/SUB, which expands as

$$\begin{aligned} R/(MAIN/SUB) &= R/(MAIN + MAIN \cdot SUB) \\ &= R + R \cdot MAIN + \underline{R \cdot MAIN \cdot SUB} \end{aligned} \quad (7)$$

A main-plot within a complete block can be identified by specifying the level of A tested on the main-plot. Also, a sub-plot within a main-plot can be identified by the level of B tested on the sub-plot. Thus, the block model can also be expressed as R/A/B, which expands as $R : R \cdot A + \underline{R \cdot A \cdot B}$. The treatment model is $A \times B = A + B + A \cdot B$, where A · B is the interaction of A and B, so the full model is as given in eqn (4).

Example 7
When many treatments are to be tested, a complete replication may be subdivided into incomplete blocks. Each treatment is tested once in a replication. This type of design is denoted as lattice design and comes in different variants (Mead 1997). The layout of a lattice design for a single treatment factor involves two *nested* randomization units: plots (PLOT) within incomplete blocks and incomplete blocks (IBLOCK) nested within complete replications (R). Thus, the block model is R/IBLOCK/PLOT, which expands as

$$R : R \cdot IBLOCK + \underline{R \cdot IBLOCK \cdot PLOT} \quad (8)$$

Example 8
Two factors A and B are laid out as a split-block (strip-plot) design with A in main-rows (MAINROW) and B in main-columns (MAINCOL). The randomization of main-rows and main-columns is *crossed*. Each plot in this design is regarded as a main-plot, which serves as a block to accommodate two additional factors C and D. These are laid out as a split-block with C in sub-rows (SUBROW) and D in sub-columns (SUBCOL). The treatment model is $A \times B \times C \times D$, while the block model is $R/(MAINROW \times MAINCOL)/(SUBROW \times SUBCOL)$. Within a block, main-rows and main-columns are uniquely identified by A and B. Similarly, sub-rows and sub-columns in a main-plot are uniquely identified by C and D. Thus, the block model can also be given as $R/(A \times B)/(C \times D)$ which reflects the fact that the

crossed randomization for C and D is *nested* within that for A and B. The block model expands as

$$R : R \cdot A + R \cdot B + R \cdot A \cdot B + R \cdot A \cdot B \cdot C + R \cdot A \cdot B \cdot D + \underline{R \cdot A \cdot B \cdot C \cdot D} \quad (9)$$

The first three random terms represent the randomization structure for A and B as a split-block. The random terms involving C and D are all nested within $R \cdot A \cdot B$.

Interaction among block and treatment factors

Rule 6 (interaction among block and treatment factors)

After setting up the block and treatment models, assuming absence of block-by-treatment interaction, one may check if there is reason to assume an interaction among a block factor and a treatment factor. The need for an interaction term among a block and a treatment factor may occur when heterogeneity among different block units is expected to be large. An interaction term can be added to the model, providing the design allows estimation of the interaction. A prerequisite for estimability of a block-by-treatment interaction is that valid analyses are possible for each level of the block factor. An analysis of block-by-treatment interaction may be of major interest, e.g. when blocks are different environments.

Example 9

Randomized complete block experiments testing different genotypes (G) are replicated across several randomly selected locations (L). Blocks are coded by R, while plots are coded by PLOT. Locations are an additional blocking factor, which is superimposed on complete blocks. The block model is given by L/R/PLOT, while the treatment model equals G. Assuming no block-by-treatment interaction, the full model is $G : L + L \cdot R + \underline{L \cdot R \cdot PLOT}$.

Differences among locations are usually so large, that a genotype-by-location interaction ($L \cdot G$) must be postulated. Estimation of $L \cdot G$ is feasible because a full analysis is possible for each location. Adding this interaction, the model is

$$G : L + L \cdot G + L \cdot R + \underline{L \cdot R \cdot PLOT} \quad (10)$$

Multi-phase experiments

Some experiments can be subdivided into several phases, with a different design used in each phase. In a two-phase design, e.g. observational units from the first phase are taken to the second phase, where they are randomized according to a new design (Brien 1983).

Rule 7 (multi-phase experiments)

If an experiment involves several phases, set up the block model for each and add block models from different phases to the overall model. Any repetitions of an effect in block models from different phases are deleted. Similarly, if several effects are confounded, only one of the confounded effects is retained. The block model of a phase pertains to observational units in that phase, i.e. all block factors identify and are innate to the observational units of that phase. Samples obtained from observational units at a phase are allocated by a randomization scheme to observational units of a subsequent phase. Any steps taken prior to randomized allocation in the subsequent phase are considered as belonging to the immediately preceding phase.

Rule 7 is essentially equivalent to the set of rules given by Brien (1983) and Brien and Payne (1999) (also see http://www.maths.unisa.edu.au/matcjb/multitier/). Appealing to Rule 6, interactions among treatment factors and block factors from different phases may be added as deemed appropriate, providing the design allows identification of these effects.

Example 10

The present example considers a variant of the split-plot design discussed by Cochran and Cox (1957, Section 7.33) and, in a slightly different form, by Hinkelmann and Kempthorne (1994, Section 13.4.3). Assume that in the first phase the treatment factor harvesting date (A) is tested in complete blocks (R). On each plot, harvested plants are bulked and subsequently split into three subsamples, each of which is to be analysed for a nutrient by one of three different chemical methods, defining a second treatment factor B with three levels. For a plot, the three methods are randomly allocated to the three samples. The design at phase I is a split-plot with block model $R/A/B = R : R \cdot A + \underline{R \cdot A \cdot B}$. The chemical analysis in the lab constitutes phase II of the experiment. Samples are dried in the lab before chemical analysis. It is a curiosity and, in fact, a weakness of this design that the treatment factor A (harvesting date) needs to be used as a blocking factor in time, because harvested plants must be processed immediately, i.e. samples

from different harvest dates cannot be dried at the same time. Clearly, main effects for harvest dates are confounded with block effects from phase II, so strictly speaking, the design does not permit inferences with respect to harvest date main effects. Three ovens (O) are available for drying. They must be used simultaneously to make subsequent chemical analyses comparable. The three ovens are just large enough for drying all plant samples of a harvest date. They differ in the speed with which the plant samples are dried. This difference in speed may affect chemical analyses. Thus, the type of oven is used as a blocking variable. Specifically, the main plot from the first phase is taken as the column of a Latin square for testing factor B, while the oven type (O) is taken to be the row (Fig. 1). A separate Latin square is used for each harvest date (A), i.e. Latin squares are nested within A. For a given level of A, the main-plot is uniquely identified by the block from phase I (R). Thus, O and R identify rows and columns of each Latin square and the block model for the second phase is $A/(R \times O) = A : A \cdot R + A \cdot O + \underline{A \cdot R \cdot O}$. Note that the main-plot effect $A \cdot R$ occurs in the block model of both phases because both phases use the main-plot as a randomization unit. In addition, the block model of phase II contains a main effect for A, which serves as a blocking variable. Only one copy of the effects A and $A \cdot R$ is retained in the final model. Moreover, both $\underline{R \cdot A \cdot B}$ and $\underline{A \cdot R \cdot O}$ are effects with a separate level for each observational unit (sample) and so both represent a residual. Clearly, the two effects are confounded, so only one of the two terms is retained. Note that it is impossible to estimate separate variance components for confounded random effects. It should be stressed, however, that the two residual effects are not identical. The term $\underline{R \cdot A \cdot B}$ represents sub-sampling error from phase I, while $\underline{A \cdot R \cdot O}$ accounts for added residual variation incurred in the drying process of phase II. Including the treatment model $A \times B$, the full model reads

$$A + B + A \cdot B + R : A \cdot O + R \cdot A + \underline{R \cdot A \cdot B} \quad (11)$$

This is essentially the model given by Cochran and Cox (1957) and by Hinkelmann and Kempthorne (1994), except that we take $A \cdot O$ random to reflect the randomization structure.

Taking random effects fixed

Rule 8 (taking random effects fixed)
A random effect, which does not *contain* (is not aliased with) the fixed effect to be tested (F-test, t-test, etc.), may be taken as fixed in the analysis. This is advantageous mainly when the random effect has less than five or 10 levels, in which case variance component estimates will be very unreliable. To decide whether a random effect contains the fixed effects to be tested, one needs to express the random effect in terms of the factors involved in the fixed effect, if possible, as described in Rule 5. Additionally, a simple check works as follows: Replace the random effect by its cross with the treatment effect to be tested. If the resulting analysis remains completely unaltered, the fixed effects is contained in the random effect, but not otherwise.

Example 7 (continued)
The model for a lattice design in eqn (8) involves a random term, $R \cdot IBLOCK$, for incomplete blocks. This term does not include the treatment effect A, as can be verified from the fact that replacing $R \cdot IBLOCK$ by $R \cdot IBLOCK \cdot A$ changes the result of the analysis. Thus, we may formally treat $R \cdot IBLOCK$ as fixed, when making statistical inferences regarding A. The resulting analysis will exploit only the intra-block information and is therefore known as intra-block analysis (Cochran and Cox 1957, Mead 1997, Federer and Wolfinger

a_1				a_2				a_3				a_4			
b_3	b_2	b_1	o_1	b_1	b_3	b_2	o_3	b_3	b_1	b_2	o_2	b_2	b_3	b_1	o_2
b_1	b_3	b_2	o_2	b_2	b_1	b_3	o_2	b_1	b_2	b_3	o_3	b_1	b_2	b_3	o_1
b_2	b_1	b_3	o_3	b_3	b_2	b_1	o_1	b_2	b_3	b_1	o_1	b_3	b_1	b_2	o_3
III	II	I		III	I	II		II	I	III		I	III	II	

Fig. 1: Phase II design (laboratory) with four Latin squares for factor B (chemical method) with levels b_1, b_2 and b_3 (Example 10). There is one Latin square for each level of A (harvesting date; a_1, a_2, a_3 and a_4). For each Latin square, rows are ovens (o_1,o_2 and o_3), while columns correspond to blocks (I, II and III) from phase I. The ovens in each Latin square are the same

1998, Federer et al. 2001). By contrast, the model with random blocks also utilizes the inter-block information. Recovery of inter-block information is worthwhile when the block variance is not large relative to the residual variance and when the number of blocks is large enough to obtain reliable estimates of the block variance component. Quite often, both analyses yield rather similar results.

Example 9 (continued)
In model (10), we may take all effects as fixed, which do not contain G. This yields

$$G + L + L \cdot R : L \cdot G + \underline{L \cdot R \cdot PLOT} \qquad (12)$$

This model will provide an analysis, in which genotype comparisons are based solely on comparisons within experiments, while the between-experiment information is not exploited. The analysis is analogous to an intra-block analysis for an incomplete block design (Patterson 1997). By contrast, analysis using (10) also exploits the inter-experiment as well as the inter-block information, which will be present when the data are unbalanced. Quite often in series of experiments for cultivar evaluation, variances associated with random terms not involving G are large, so the between-experiment information is small and both analyses yield similar results. Analysis by model (12) is useful mainly when the number of locations is small.

Statistical Inference

It is not our aim to provide an extensive description of the theoretical underpinnings of statistical inference for mixed models, which can be found elsewhere (e.g. Hocking 1985, Searle et al. 1992, Verbeke and Molenberghs 2000). Our objective is to briefly discuss a few issues that are relevant for users of mixed model packages. In particular, we will contrast classical ANOVA procedures to likelihood-based mixed model analyses now in common use.

Missing data

Quite often, designed experiments give rise to missing values due to unforeseen circumstances, so that the data set available for analysis is smaller than planned by design. Fortunately, statistical inference for mixed models remains valid, providing the data meet the *missing completely at random* (MCAR) assumption (Verbeke and Molenberghs 1997). Loosely speaking, the MCAR assumption requires that the missing data pattern is independent of the design, particularly the treatment structure. A more rigid definition of the MACR assumption is given, e.g., in Verbeke and Molenberghs (1997). In fact, statistical inference remains valid under somewhat milder conditions than the MCAR assumption, as discussed in Verbeke and Molenberghs (1997). Care is needed in deciding whether or not the MCAR assumption is satisfied, as shown in the next two examples.

Example 11
Three sub-samples per plant are analysed in the lab for a plant nutrient. A test tube is accidentally dropped in the lab, thus giving rise to missing data. Which particular test tube is dropped will not usually be influenced by the treatment or randomization units corresponding to the test tube, so data are missing completely at random.

Example 12
A pot experiment is performed to evaluate three different fertilizers. Each pot contains ten plants at the onset of the experiment. One of the fertilizers has a harmful effect on some plants, causing them to die off during the experiment. The resulting missing data pattern does not meet the MCAR assumption – which particular plants are missing depends on the treatment.

Ordinary least squares vs. generalized least squares

Fixed effects parameters of a linear model may be estimated by the method of ordinary least squares (OLSE). This method is optimal in linear models with a single error term and homogeneous variances (Searle et al. 1992). In mixed linear models, however, OLSE is not usually optimal. A better method is known as generalized least squares (GLSE). This method uses weights computed from the variance components of random effects. GLSE has optimal properties when the variance components are known. Parameter estimators then are best linear unbiased estimators, i.e. they have minimum variance among all unbiased linear estimators. In practice, variance components need to be estimated, but GLSE is usually better than OLSE, even if variance component estimates are used. OLSE is the estimation method used in the SAS procedure GLM, while the MIXED procedure employs GLSE. Parameter estimates by the two

methods do not usually agree, except for some special cases (Searle et al. 1992), e.g. when the data are balanced (Piepho and Spilke 1999).

Estimation of variance components

There are several methods for the estimation of variance components (Searle et al. 1992). Restricted maximum likelihood (REML) (Patterson and Thompson 1971) has come to be the method of choice, and it is the default in most mixed model packages. The output of REML-based mixed model procedures usually contains a number of descriptive measures related to the maximized restricted likelihood (or log-likelihood), which will be unfamiliar to those accustomed mainly to classical ANOVA procedures. These statistics may be used for assessing model fit and for selecting among several candidate models, as described by Wolfinger (1996), and are relevant mainly when using mixed models, which are not of the simple ANOVA type.

When a random term has few levels, the variance component estimate may become unreliable. In this case, it should be checked whether the effect can be taken as fixed by Rule 8. Also, for balanced data, it is worthwhile to drop the non-negativity constraint on variance component estimates (NOBOUND option in PROC MIXED of SAS). This will make REML estimates identical with ANOVA estimates, which are optimal for balanced data and yield exact Wald-tests for fixed effects (Piepho and Spilke 1999).

ANOVA vs. REML-based Wald-tests

Tests of fixed effects in a mixed model may be performed by two different strategies, which do not necessarily yield identical results. The first is based on a classical decomposition of the total sum of squares into components attributable to different model effects. F-tests are constructed based on the expected values of ANOVA mean squares, which can be computed using a general algorithm described, e.g. by Milliken and Johnson (1984). Specifically, F-statistics are constructed as a ratio of linear combinations of mean squares so that the numerator and denominator expected values are the same except for an additional term in the numerator depending on the effect to be tested. This method is implemented, e.g., in the GLM procedure of SAS. It is preferable to more *ad hoc* methods such as that proposed by Heyland and Kochs (1978). In case the numerator or the denominator of the F-statistic involve more than one mean square, the Satterthwaite-method (Milliken and Johnson 1984) may be used to compute degrees of freedom.

The second method computes Wald-type F-statistics using GLSE of fixed effects based on REML estimates of the variance components (Littell et al. 1996). It is the method employed by the MIXED procedure of the SAS system. The printed output of MIXED will be familiar to those accustomed to classical ANOVA in that an F-test is produced for each fixed effect in the model. What is unfamiliar is the lack of sums of squares and mean squares. The distribution of the F-statistic under the null hypothesis usually needs to be approximated. The currently best approximation method is that proposed by Kenward and Roger (1997), which is available via the DDFM option of the MIXED procedure.

For balanced data, ANOVA F-tests and Wald-tests produce identical results, providing REML is not constrained to produce non-negative variance estimates (Piepho and Spilke 1999), but results for unbalanced data differ. It is not clear, which of the two methods is preferable. Wald-tests use GLSE, which is optimal only when variance components are known, as is rarely the case in practice.

In the analysis of linear models with a single error term, sums of squares may be computed by different methods. The most common methods are sequential sums of squares, denoted as Type I SS in SAS and Type III SS (Searle 1987). These same methods are used for a mixed model ANOVA. In addition, Wald-type F-tests may be defined for testing Type I or Type III hypothesis. Nelder (1994) has a lucid discussion of Type I and Type III hypothesis tests, and provides strong arguments in favour of Type I hypothesis testing, e.g. better power. Type III hypotheses are relevant mainly if one is prepared to test main effects for a fixed factor when interactions with another fixed factor are significant. The Type III vs. Type I controversy essentially becomes a non-issue, if one adheres to the philosophy of testing main effects and marginal means only when the interaction is not significant. For details see Nelder (1994).

Mean comparisons

Treatment means may be computed from least squares estimates of fixed effects (either OLSE or GLSE). These means are known as least square

means (LS-means). When the data are balanced, the standard error of a difference (S.E.D.) among LS-means is the same for each pairwise comparison. In this case, a common critical difference may be computed. For example, Fisher's least significant difference is given by $LSD = t \times S.E.D.$, where t is a critical value from a t-distribution with appropriate degrees of freedom. In addition, a lines display or letters display can be obtained by standard procedures (Steel and Torrie 1980).

When data are unbalanced, the S.E.D. is not constant among comparisons, and hence there is no common critical difference. In addition, a lines display is not forthcoming by standard approaches. This is disappointing for those used to lines displays and critical differences. Some packages, e.g. GENSTAT and ASREML, report an average S.E.D. across all pairwise comparisons, which may be reported in place of a critical difference as a measure of accuracy. The MIXED procedure of SAS produces all pairwise comparisons in a line-by-line fashion, together with an S.E.D. for each comparison, but no average S.E.D. (this can be computed in a subsequent datastep). H. P. Piepho (2003) has developed an algorithm which produces a letters display for unbalanced as well as for balanced data. An SAS macro is available from the author upon request.

The GLM procedure of SAS computes LS-means based on OLSE and it does not compute appropriate S.E.D. Thus, the GLM procedure should not be used for mean comparisons. The MIXED procedure uses GLSE to obtain LS-means. Standard errors are computed from REML estimates of the variance components, and degrees of freedom may be derived using the method of Kenward and Roger (1997).

Multiple testing entails the danger of an inflated Type I error rate. For balanced data, many specialized methods are available for controlling the family-wise Type I error rate (FWE), e.g. the Tukey procedure (Hsu 1996). These methods are not applicable in the unbalanced mixed linear model. A general-purpose method for FWE control is the simulation approach by Edwards and Berry (1987), which is implemented in the MIXED procedure. A number of alternative methods are discussed by Westfall et al. (1999), who provide several SAS macros.

Prediction of random effects (BLUP)

There is an important exception to Rule 1. A treatment factor may be considered random due to the sampling design, and yet there is an interest in the specific treatment levels tested in an experiment. If the number of levels is large, it is advantageous to consider the factor as random and obtain estimates of random effects under that model (Searle et al. 1992, p. 18), e.g. in a plant breeding trial evaluating a large set of lines derived from a single cross. A popular estimation method for this purpose is known as best linear unbiased prediction (BLUP) and it is often used in plant and animal breeding (Searle et al. 1992, Mrode 1998). BLUPs may be more efficient than estimators assuming fixed effects. In most agronomic trials, however, the treatment levels are purposefully selected and the number of levels is small, so the assumption of random sampling is not tenable.

A Brief Review of Some Extensions to ANOVA-Type Mixed Models as Relevant in Agronomy

This paper has focused on ANOVA-type mixed models. There are a number of extensions to this type of mixed model, which will be briefly reviewed in this section. A full coverage is beyond the scope of this paper, and the reader interested in more details is referred to the pertinent literature.

Often, repeated measurements are taken on the same experimental unit, the repetition being either in space or in time or both. Repeated measurements call for special attention due to the correlation induced by the fact that the same experimental unit is involved in several non-randomized measurements.

Example 13

A randomized complete block experiment is conducted to evaluate four different fertilizers for a perennial grass crop. On each plot, repeated measurements of yield are made in successive years, which define a second (repeated) treatment factor. Observations made on the same plot in successive years are correlated. The correlation among neighbouring years can be expected to be higher than, e.g. that among the first and the last.

The analysis of repeated measurements needs to account for the correlation among observations on the same experimental unit. There are many different types of model that can be imposed on the correlation structure, including autoregressive and spatial models. These correlation structures can be embedded in a mixed model framework, but the resulting models are no longer of the

ANOVA-type (Davidian and Giltinan 1995, Diggle et al. 1996, Hand and Crowder 1996, Verbeke and Molenberghs 1997, 2000, Schabenberger and Pierce 2001). Good related key words for searching the literature are *spatial statistics*, *geostatistics*, *time series* and *longitudinal data*. Long-term rotation experiments pose a number of specific repeated-measures problems, which require specialized modelling approaches (Singh et al. 1997).

A simple method for the analysis of repeated measurements is to compute a summary statistic per the smallest randomized observational unit (plot, pot, etc.). This will yield one data point per unit, so the data can, in fact, be analysed by ANOVA-type mixed models (Diggle et al. 1996). For example, phytopathologists use the 'area under the curve' as an integral measure of disease incidence over a period of time (Campbell and Madden 1990). Trends in time or space across a short series of repeated measurements on the same unit can often be assessed by simple linear regression, yielding one slope estimate per experimental unit (Bürkert et al. 2002) or by non-linear regression parameters. Often, a simple mean across the series is a useful summary, e.g. in a multi-year grazing trial.

Many authors have suggested the analysis of repeated measurements as if the repeated treatment factor were the split-plot factor in a split-plot design ('split-plot in time'; Steel and Torrie 1980, Peterson 1994). This approach may be justly criticized, mainly on the grounds that a justification by randomization theory is not forthcoming (Hinkelmann and Kempthorne 1994). Its popularity stems from the fact that analysis is relatively simple. With the advent of powerful computers and advanced mixed model software, the split-plot approach has become largely obsolete. A split-plot model implies a very simple correlation structure for repeated measurements, and strictly speaking its use is justified only when a better fit compared with more complex correlation structures has been established (Wolfinger 1996).

The mixed model approach to repeated measurements needs to be contrasted with multivariate ANOVA methods (MANOVA) (Cole and Grizzle 1966). The MANOVA approach has two important drawbacks: (i) The sample size required for multivariate test statistics to be computable is usually larger than for test statistics under a mixed model. In fact, the necessary sample size may be prohibitive for agronomic trials. (ii) The MANOVA approach can analyse only complete series of repeated measurements. In Example 13, if one of the measurements on the same plot is missing, all measurements on that plot must be omitted from the analysis. By contrast, with mixed models there is no need to omit any observations. The main advantage of the MANOVA approach lies in the less restrictive statistical assumptions compared with mixed modelling. For more details see Hand and Crowder (1996) and Diggle et al. (1996).

Large field trials pose formidable problems in terms of error control and classical blocked designs are not always effective. An alternative approach is to use spatial models for the correlation among neighbouring plots. Many of the methods of spatial analysis for field trials can be embedded in the mixed model framework (Cullis and Gleeson 1991, Gilmour et al. 1997, Gleeson 1997, Wu et al. 1998). Spatial models may also be combined with smoothing splines to account for large-scale field trends (Verbyla et al. 1999).

In this paper, we basically neglected the treatment structure and mainly looked at the block model, as this may involve several random effects, thus rendering the linear model of the mixed type. For simplicity, all treatment factors were assumed to be qualitative. If a treatment factor is quantitative such as fertilizers or pesticides tested in different quantities, one may consider a regression. Regression in mixed models falls in the ANOVA framework so long as all regression coefficients are fixed. When some coefficients are random, one may need to use more refined modelling (Wolfinger 1996, Verbeke and Molenberghs 1997).

The analysis of a series of experiments is complicated due to the presence of treatment-by-environment interaction. ANOVA-type mixed models cannot always satisfactorily model interaction and several extensions have been suggested, including heteroscedastic (Denis et al. 1997, Frensham et al. 1997, Piepho 1999a) and multiplicative (Piepho 1997, Smith et al. 2001) models. A topic related to a series of experiments is stability analysis. Measures of stability are very popular among plant breeders and agronomists as a means to assess yield variability across varying environments. Many stability measures can be expressed as functions of parameters of a mixed model (Piepho 1998, 1999a), and so mixed models play a key role in stability analysis. In addition, spatial models at the field trial level can be integrated with advanced models for treatment-by-environment interaction within a single mixed model (Smith et al. 2001).

Mixed models have also been extended to allow for non-normal data. The extension is known as the

generalized linear mixed model (Schabenberger and Pierce 2001, McCulloch and Searle 2001), and it is highly relevant for problem data such as count data (e.g. weeds and insects) and percentages (e.g. disease incidence, weed coverage and emergence rates) (Piepho 1999b). Moreover, non-linear mixed models can accommodate intrinsically non-linear regression models (Davidian and Giltinan 1995), e.g. plant growth models (Gregoire and Schabenberger 1996, Schabenberger and Pierce 2001).

Acknowledgements

The authors are grateful to Hanspeter Thöni (Universität Hohenheim, Germany) and Christel Richter (Humboldt-Universität Berlin, Germany) for helpful comments on earlier versions of the manuscript. Thanks are also due to Brian Cullis for drawing attention to the important work of Chris Brien.

References

Bailey, R. A., 1991: Strata for randomized experiments (with discussion). J. R. Stat. Soc. B **53,** 27—78.

Brien, C. J., 1983: Analysis of variance tables based on experimental structure. Biometrics **39,** 53—59.

Brien, C. J., and R. W. Payne, 1999: Tiers, structure formulae and the analysis of complicated experiments. The Statistician **48,** 41—52.

Bürkert, A., H. P. Piepho, and A. Batino, 2002: Multi-site time trend analysis of crop residue, phosphorus, nitrogen and legume rotation effects on cereal yields in sub-Saharan West Africa. Exp. Agric. **38,** 163—183.

Campbell, C. L., and L. V. Madden, 1990: Introduction to Plant Disease Epidemiology. Wiley, New York.

Cochran, W. G., and G. M. Cox, 1957: Experimental designs. Wiley, New York.

Cole, J. W. L., and J. E. Grizzle, 1966: Applications of Multivariate Analysis of Variance to Repeated Measures Experiments. Biometrics **22,** 810—828.

Cullis, B. R., and A. C. Gleeson, 1991: Spatial analysis of field experiments – an extension to two dimensions. Biometrics **47,** 1449—1460.

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

Denis, J.-B., H. P. Piepho, and F. A. van Eeuwijk, 1997: Modelling expectation and variance for genotype by environment data. Heredity **79,** 162—171.

Diggle, P. J., K.-Y. Liang, and S. L. Zeger, 1996: Analysis of Longitudinal Data. Clarendon Press, London.

Edwards, D., and J. J. Berry, 1987: The efficiency of simulation-based multiple comparisons. Biometrics **43,** 913—928.

Federer, W. T., and R. D. Wolfinger, 1998: SAS code for recovering intereffect information in experiments with incomplete block and lattice rectangle designs. Agron. J. **90,** 545—551.

Federer, W. T., M. Reynolds, and J. Crossa, 2001: Combining results from augmented designs over sites. Agron. J. **93,** 389—395.

Frensham, A., B. R. Cullis, and A. P. Verbyla, 1997: Genotype by environment variance heterogeneity in a two-stage analysis. Biometrics **53,** 1373—1383.

Gilmour, A. R., B. R. Cullis, and A. P. Verbyla, 1997: Accounting for natural and extraneous variation in the analysis of field experiments. J. Agric. Biol. Environ. Stat. **2,** 269—293.

Gleeson, A. C., 1997: Spatial analysis. In: R. A. Kempton, and P. N. Fox (eds), Statistical Methods for Plant Variety Evaluation, pp. 68—85. Chapman & Hall, London.

Gregoire, T. G., and O. Schabenberger, 1996: Nonlinear mixed-effects modeling of cumulative bole volume with spatially correlated within-tree data. J. Agric. Biol. Environ. Stat. **1,** 107—119.

Hand, D. J., and M. Crowder, 1996: Practical Longitudinal Data Analysis. Chapman & Hall, London.

Heyland, K.-U., and H.-J. Kochs, 1978: Varianzanalytische Testung mehrfaktorieller pflanzenbaulicher Versuche. Zeitschrift für Acker und Pflanzenbau **146,** 109—119.

Hinkelmann, K., and O. Kempthorne, 1994: Design and analysis of experiments. Introduction to Experimental Design, Vol. 1. Wiley, New York.

Hocking, R. R., 1985: The Analysis of Linear models. Brooks and Cole, Monterey.

Hsu, J. C., 1996: Multiple Comparisons. Chapman & Hall, London.

Kenward, M. G., and J. H. Roger, 1997: Small sample inference for fixed effects from restricted maximum likelihood. Biometrics **53,** 983—997.

Littell, R. C., G. A. Milliken, W. W. Stroup, and R. D. Wolfinger, 1996: SAS System for Mixed Models. SAS Institute, Cary, NC.

McCullagh, P., and J. A. Nelder, 1989: Generalized Linear Models, 2nd edn. Chapman & Hall, London.

McCulloch, C. E., and S. R. Searle, 2001: Generalized, Linear, and Mixed Models. Wiley, New York.

McLean, R. A., W. L. Sanders, and W. W. Stroup, 1991: A unified approach to mixed linear models. The Am. Stat. **45,** 54—64.

Mead, R., 1997: Design of plant breeding trials. In: R. A. Kempton, and P. N. Fox (eds), Statistical Methods for Plant Variety Evaluation, pp. 40—67. Chapman & Hall, London.

Milliken, G. A., and D. E. Johnson, 1984: Analysis of Messy Data. Designed Experiments, Vol. 1. Chapman & Hall, London.

Mrode, R. A., 1998: Linear Models for the Prediction of Animal Breeding Values. CAB International, Wallingford.

Nelder, J. A., 1965: The analysis of randomized experiments with orthogonal block structure. I. Block structure and the null analysis of variance. II.

Treatment structure and the general analysis of variance. Proc. R. Soc. Lond. A **283,** 147—178.
Nelder, J. A., 1994: The statistics of linear models: back to basics. Stat. Comput. **4,** 221—234.
Patterson, H. D., 1997: Analysis of series of variety trials. In: R. A. Kempton, and P. N. Fox (eds), Statistical Methods for Plant Variety Evaluation, pp. 139—161. Chapman & Hall, London.
Patterson, H. D., and R. Thompson, 1971: Recovery of inter-block information when block sizes are unequal. Biometrika **58,** 545—554.
Payne, R. W., and G. N. Wilkinson, 1977: A general algorithm for analysis of variance. Appl. Stat. **26,** 251—260.
Peterson, R. G., 1994: Agricultural Field Experiments. Design and Analysis. Marcel Dekker, New York.
Piepho, H. P., 1997: Analyzing genotype-environment data by mixed models with multiplicative effects. Biometrics **53,** 761—766.
Piepho, H. P., 1998: Methods for comparing the yield stability of cropping systems – a review. J. Agron. Crop Sci. **180,** 193—213.
Piepho, H. P., 1999a: Stability analysis using the SAS system. Agron. J. **91,** 154—160.
Piepho, H. P., 1999b: Analysing disease incidence data from designed experiments by generalized linear mixed models. Plant Pathol. **48,** 668—674.
Piepho, H. P., 2003: An algorithm for a letter-based representation of all-pairwise comparisons. Journal of Computational and Graphical Statistics (in press).
Piepho, H. P., and J. Spilke, 1999: Anmerkungen zur Analyse balancierter gemischter Modelle mit der SAS Prozedur MIXED. Zeitschrift für Agrarinformatik **7,** 39—46.
Schabenberger, O., and F. J. Pierce, 2001: Contemporary Statistical Models. CRC Press, Boca Raton.
Searle, S. R., 1987: Linear Models for Unbalanced Data. Wiley, New York.
Searle, S. R., G. Casella, and C. E. McCulloch, 1992. Variance Components. Wiley, New York.
Singh, M., S. Christiansen, and B. K. Chakraborty, 1997: An assessment of the effect of covariances of plot errors over time on the precision of means of rotation experiments. Exp. Agric. **33,** 469—475.
Smith, A., B. R. Cullis, and R. Thompson, 2001: Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. Biometrics **57,** 1138—1147.
Steel, R. G. D., and J. H. Torrie, 1980: Principles and Procedures of Statistics: a Biometrical Approach, 2nd edn. McGraw-Hill, New York.
Verbeke, G., and G. Molenberghs (eds), 1997: Linear mixed Models in Practice. A SAS-oriented Approach. Springer, Berlin.
Verbeke, G., and G. Molenberghs, 2000: Linear Mixed Models for Longitudinal Data. Springer, Berlin.
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. Appl. Stat. **48,** 269—300.
Westfall, P. H., R. D. Tobias, D. Rom, R. D. Wolfinger, and Y. Hochberg, 1999: Multiple Comparisons and Multiple Tests. SAS Institute, Cary, NC.
Wilkinson, G. N., and C. E. Rogers, 1973: Symbolic description of factorial models for analysis of variance. Appl. Stat. **22,** 392—399.
Wolfinger, R. D., 1996: Heterogeneous variance-covariance structures for repeated measures. J. Agric. Biol. Environ. Stat. **1,** 205—230.
Wu, T. X., D. E. Mather, and P. Dutilleul, 1998: Application of geostatistical and neighbor analyses to data from plant breeding trials. Crop Sci. **38,** 1545—1553.