

REGULAR ARTICLE

A two-stage approach to recovery of inter-block information and shrinkage of block effect estimates

Hans-Peter Piepho ^{1*}, Emlyn R. Williams ², Joseph O. Ogutu ¹

¹ Bioinformatics Unit, Institute of Crop Science, University of Hohenheim, Fruwirthstrasse 23, 70599 Stuttgart, Germany.

² Statistical Consulting Unit, The Australian National University, Canberra ACT 0200, Australia.

*Corresponding author: Hans-Peter Piepho; E-mail: piepho@uni-hohenheim.de

CITATION: Piepho, H.P., Williams, E.R., Ogutu, J. (2013). A two-stage approach to recovery of inter-block information and shrinkage of block effect estimates. *Communications in Biometry and Crop Science* 8 (1), 10–22.

Received: 14 January 2013, Accepted: 27 May 2013, Published online: 25 June 2013
© CBCS 2013

ABSTRACT

Linear mixed models often comprise several effects, and the focus is usually only on one or a few of them, while the other effects need to be fitted merely to adjust for all sources of variation. A typical example is the analysis of a blocked experiment, where the effects of interest pertain to treatments, while effects for replicates and incomplete blocks need to be taken into account in order to obtain efficient estimates of treatment effects. If fixed and random effects that are not of major interest were known beforehand, we could subtract these from the observed data, and a reduced model could then be fitted to these corrected data in order to estimate the effects of interest. In practice, this approach cannot be used directly, however, because true values of the effects are unknown. But we may replace unknown effects by their estimates. We show in this paper, that a reduced model fitted to these 'empirically' corrected data yields BLUE and BLUP of fixed and random effects of interest in the full model. Using examples, this result is demonstrated to be useful for illustrating the recovery of inter-block information and for understanding the properties of estimators obtained from mixed-model analysis.

Key Words: *intra-block analysis; linear model; linear mixed model; recovery of inter-block information; shrinkage; sweeping.*

INTRODUCTION

Linear models can be fitted to data in order to make inferences for effects of interest, while controlling for other (nuisance) effects that are not themselves of primary interest, but are important sources of variation. For example, analysis of data from a blocked experiment can be based on the linear model

$$y_{ij} = \mu_i + b_j + e_{ij}, \quad (1)$$

where y_{ij} is the response of the i -th treatment ($i = 1, \dots, I$) in the j -th block ($j = 1, \dots, J$), μ_i is the mean of the i -th treatment, b_j is the fixed effect of the j -th block, subject to the constraint $\sum_{j=1}^J b_j = 0$, and e_{ij} is the random error term associated with y_{ij} , assumed to be independently normally distributed with constant variance σ_e^2 . We are interested in making inferences regarding the treatment means, while block effects need to be accounted for to represent the randomization layout and to increase precision.

Analysis based on model (1) entails estimating all effects simultaneously and then deriving inference for effects of interest. In our experience from teaching linear model concepts, however, many researchers prefer a two-stage view, according to which the data are first corrected for nuisance effects, and the corrected data, now presumed to be free of nuisance effects, are then subjected to analysis focusing on the effects of interest. In the example, if we knew the values of block effects, we could subtract these from both sides of (1) and then perform an analysis based on the resulting reduced model, which in this case would boil down to a one-way analysis of variance. In practice, the nuisance effects are not known but have to be estimated. The purpose of this paper is to show that an analysis of the corrected data using a model reduced by first accounting for estimated nuisance effects yields the best point estimates of fixed and random effects of interest, i.e., the best linear unbiased estimators (BLUE) and best linear unbiased predictors (BLUP), respectively. The method does not usually provide any computational advantages over the standard procedures, but we have found it to be useful for teaching purposes because of its intuitive appeal and interpretive insights. In particular, the method can be used to explain the so-called recovery of inter-block information (Yates, 1940) in blocked experiments. In this paper, we first develop the general result in Section 2. This is then applied to special cases in Section 3.

THE LINEAR MIXED MODEL

Consider the linear mixed model (McLean et al., 1991; Searle et al., 1992)

$$y = X\beta + Zu + e, \quad (2)$$

where X and Z are known design matrices for fixed effects β and random effects u , respectively. The expected values of random effects are $E(u) = 0$ and $E(e) = 0$, the variances are $\text{var}(u) = G$ and $\text{var}(e) = \sigma_e^2 I$, and the covariance is $\text{cov}(u, e) = 0$, so that $\text{var}(y) = V = ZGZ^T + \sigma_e^2 I$. It is assumed here that the positive-definite variance-covariance matrix G and the residual variance σ_e^2 are known. We further assume that model (2) can be partitioned as follows:

$$X\beta = X_1\beta_1 + X_2\beta_2 \quad \text{and} \quad (3)$$

$$Zu = Z_1u_1 + Z_2u_2, \quad (4)$$

where $\text{var}(u_1) = G_1$, $\text{var}(u_2) = G_2$ and $\text{cov}(u_1, u_2) = 0$. We assume here that both X_1 and X_2 have full column rank and that $\text{sp}(X) = \text{sp}(X_1) \cup \text{sp}(X_2)$, where $\text{sp}(A)$ denotes the column space generated by a matrix A . The effects of principal interest are taken to be β_1 and u_1 , but the analysis also needs to account for the other effects β_2 and u_2 , which are

themselves not of intrinsic interest. The solution of the mixed model equations pertaining to β is

$$\hat{\beta} = (X^T V^{-1} X)^{-1} X^T V^{-1} y = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} C_{11} X_1^T + C_{12} X_2^T \\ C_{21} X_1^T + C_{22} X_2^T \end{pmatrix} V^{-1} y, \quad (5)$$

where we have used $(X^T V^{-1} X)^{-1} = \begin{pmatrix} X_1^T V^{-1} X_1 & X_1^T V^{-1} X_2 \\ X_2^T V^{-1} X_1 & X_2^T V^{-1} X_2 \end{pmatrix}^{-1} = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$ with A^{-1}

denoting the inverse of A . The best linear unbiased estimator (BLUE) of estimable functions $K\beta$ is given by $K\hat{\beta}$ (McLean et al., 1991). The best linear unbiased predictor (BLUP) of u is

$$\hat{u} = GZ^T V^{-1} (y - X\hat{\beta}) = \begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \end{pmatrix} = \begin{pmatrix} G_1 Z_1^T \\ G_2 Z_2^T \end{pmatrix} V^{-1} (y - X\hat{\beta}). \quad (6)$$

We now consider a two-stage representation of the BLUE and BLUP in (5) and (6), respectively, in which we correct the observed data for nuisance effects β_2 and u_2 in the first stage. In the second stage the corrected data are used to estimate β_1 and u_1 . The intuition underlying our approach is that if we knew β_2 and u_2 , we could compute the corrected data

$$y_c^* = y - X_2 \beta_2 - Z_2 u_2 \quad (7)$$

and fit the reduced model

$$y_c^* = X_1 \beta_1 + Z_1 u_1 + e \quad (8)$$

with $\text{var}(y_c^*) = V_c = Z_1 G_1 Z_1^T + \sigma_e^2 I$. In practice, we need to replace the unknown effects β_2 and u_2 with their estimators, so the corrected data are computed as

$$y_c = y - X_2 \hat{\beta}_2 - Z_2 \hat{u}_2, \quad (9)$$

where the estimates $\hat{\beta}_2$ and \hat{u}_2 are obtained from (5) and (6), respectively.

Proposition 1: One may analyse y_c in (9) as if its variance-covariance structure were the same as that of y_c^* in (8) under the reduced model. Thus, we may fit the model

$$y_c = X_1 \beta_1 + Z_1 u_1 + e_c, \quad (10)$$

where e_c is an error term associated with y_c , as if $e_c = e$. The estimators of β_1 and u_1 under this naïve model yield the BLUE of β_1 and the BLUP of u_1 under the full model (1), i.e.,

$$\tilde{\beta}_1 = (X_1^T V_c^{-1} X_1)^{-1} X_1^T V_c^{-1} y_c = \hat{\beta}_1 \quad \text{and} \quad (11)$$

$$\tilde{u}_1 = G_1 Z_1^T V_c^{-1} (y_c - X_1 \tilde{\beta}_1) = \hat{u}_1. \quad (12)$$

Thus, the single-stage estimators of β_1 and u_1 can be represented in two stages, i.e., we first compute corrected data y_c , removing the estimates of the nuisance effects β_2 and u_2 , and then use the estimators in (11) and (12). A proof of Proposition 1 is given in the Appendix.

We note here that while equations (11) and (12) yield correct point estimates of effects, they do not lend themselves for further statistical inference (significance tests or confidence intervals). For example, the variance of $\tilde{\beta}_1$ is certainly not generally equal to $(X_1^T V_c^{-1} X_1)^{-1}$. Also, a valid estimate of V_c , i.e., an estimate that is equivalent to the restricted maximum log-likelihood (REML) estimate based on the full model (2), cannot be obtained in the second stage using standard procedures, because of the approximate nature of the error model. Thus, an estimate of V_c must be available from the first stage. Moreover, in order to remove β_2 and u_2 , we estimate the full model, which already yields direct estimates of β_1 and u_1 , so our method does not provide any computational advantages. Thus, we do not suggest using this two-stage scheme in practice. The important point to note here is that we may consider this two-stage scheme not only because it leads to the known best estimators, but also because it provides some useful insights into the properties of treatment effect estimators in blocked experiments that are not readily apparent from other methods. This will be illustrated in the following examples.

SPECIAL CASES AND EXAMPLES

RESULT APPLIED TO A FIXED EFFECTS MODEL

Consider the fixed effects model

$$y = X_1 \beta_1 + X_2 \beta_2 + e, \quad (13)$$

where the notation follows that used in (1). For example, β_1 may comprise fixed treatment effects and β_2 may comprise fixed block effects in a randomized block experiment (see eq. 1). We now use corrected data

$$y_c = y - X_2 \hat{\beta}_2, \quad (14)$$

where $\hat{\beta}_2$ is the ordinary least squares estimator of β_2 under model (13). Of course, the distributional properties of y_c in (14) are not identical to those of $y - X_2 \beta_2$, but from Proposition 1, an analysis that naïvely pretends that they are, does, in fact, yield the BLUE of β_1 . Hence, using the corrected data y_c in (14), we fit the linear model

$$y_c = X_1 \beta_1 + e_c \quad (15)$$

by ordinary least squares, yielding the estimator

$$\tilde{\beta}_1 = (X_1^T X_1)^{-1} X_1^T y_c. \quad (16)$$

It follows from Proposition 1 that the two-stage estimator in (16) is equivalent to the BLUE of β_1 based on (13).

Table 1: Yoghurt tasting raw data (Mead et al., 1993, p.81) and data corrected for estimated block effects assuming either fixed or random block effects (see Table 2), $y_{c(ij)} = y_{ij} - \hat{b}_j$, and arithmetic treatment means computed from these corrected data. Fixed effects were estimated by BLUE, random effect estimates were obtained by BLUP.

Yoghurt	Data	Taster				Mean
		1	2	3	4	
A	Raw data	8	5	9		7.333
	Corrected - blocks fixed	6.625	6.750	7.500		6.958
	Corrected - blocks random	6.750	6.591	7.636		6.992
B	Raw data	6	2		2	3.333
	Corrected - blocks fixed	4.625	3.750		3.125	3.833
	Corrected - blocks random	4.750	3.591		3.023	3.788
C	Raw data	4		4	3	3.667
	Corrected - blocks fixed	2.625		2.500	4.125	3.083
	Corrected - blocks random	2.750		2.636	4.023	3.136
D	Raw data		4	7	4	5.000
	Corrected - blocks fixed		5.750	5.500	5.125	5.458
	Corrected - blocks random		5.591	5.636	5.023	5.417

Example 1: In order to illustrate our result as applied to a fixed-effects model, we consider the yoghurt tasting data described in Mead et al. (1993, p.81). The data is reproduced in Table 1. The experiment was laid out according to a balanced incomplete block design. We use model (1) for analysis, taking tasters as blocks and yoghurts as treatments. Treatment means correspond to β_1 , while block effects relate to β_2 in model (13). Least squares estimates for taster effects are given in Table 2. Note that the design matrix X_2 for fixed block effects has full column rank if we replace the last block effect by $b_J = -\sum_{j=1}^{J-1} b_j$, in

accordance with the constraint $\sum_{j=1}^J b_j = 0$, so that X_2 has $(J-1)$ linearly independent

columns. The observed data were corrected using $y_{c(ij)} = y_{ij} - \hat{b}_j$ (Table 1). We fitted the reduced model $y_{c(ij)} = \mu_i + e_{c(ij)}$ to the corrected data, which boils down to computing the arithmetic means (i.e., simple averages) of the corrected data for each treatment. These arithmetic means coincide with the adjusted treatment means obtained by directly fitting model (1), which confirms the general result in Proposition 1. This example lucidly illustrates, in an intuitive way, the fact that fitting model (1) provides an adjustment for block effects.

Table 2: Estimates of taster effects b_j obtained from a linear model package. Fixed effects were estimated by BLUE (subject to the constraint $\sum_{j=1}^J b_j = 0$), random effects were estimated by BLUP.

Effect estimate	Taster			
	1	2	3	4
BLUE	1.375	-1.750	1.500	-1.125
BLUP	1.250	-1.591	1.364	-1.023

RESULT APPLIED TO A MIXED EFFECTS MODEL

Now consider the mixed model (2) and assume that we wish to estimate fixed effects β , correcting for random effects u . In analogy to the fixed effects case, we will use corrected data

$$y_c = y - Z\hat{u}, \quad (17)$$

where \hat{u} denotes the BLUP of u based on (2), and fit the following model to the corrected data:

$$y_c = X\beta + e_c. \quad (18)$$

This leads to the estimator

$$\tilde{\beta} = (X^T X)^{-1} X^T y_c. \quad (19)$$

From Proposition 1 this estimator is equivalent to the BLUE of β under the mixed model (2).

Example 1 (continued): This time, we fit blocks (tasters) in model (1) as random effects in order to recover the inter-block information (Patterson and Thompson, 1971). Thus, treatment means pertain to β , while block effects are collected in u of (2). Note that the fixed-effects constraint $\sum_{j=1}^J b_j = 0$ is now replaced by the assumption $E(b_j) = 0$ for the random block effects. The residual maximum likelihood (REML) estimates of the variance components for blocks and error are $\hat{\sigma}_b^2 = 2.5667$ and $\hat{\sigma}_e^2 = 0.6833$, respectively. We subtract BLUPs of block effects (Table 2) from the observed response, yielding corrected data $y_{c(ij)} = y_{ij} - \hat{b}_j$, and then fit a fixed-effects model with effects for treatment means ($y_{c(ij)} = \mu_i + e_{c(ij)}$) to the corrected data (Table 1). This is equivalent to computing arithmetic means for $y_{c(ij)}$ per treatment and yields treatment means that are identical to BLUE under the mixed model.

It is noteworthy that the adjusted means based on adjustments with random block effects are closer to the arithmetic yoghurt means than are the adjusted means obtained with fixed block effects (Table 1). Our approach nicely illustrates that this observation, which generally holds for analyses of incomplete block designs, can be explained by the shrinkage of block effect estimates under the random block effects model. The fact that this shrinkage must reduce the adjustment compared to the fixed block effects model can be seen directly from the equation for corrected data ($y_{c(ij)}$), and it is readily verified by comparing the adjustments in Table 2.

RESULT APPLIED TO CORRECTION FOR BOTH FIXED AND RANDOM EFFECTS

In Example 1, we considered correction for a fixed or a random nuisance effect, when the effect of interest is fixed. In more complex models, there may be a need to account for both fixed and random effects. For example, in a resolvable row-column design, rows and columns are incomplete blocking units nested within complete replicates. In the analysis, one may want to correct for fixed effects of complete replicates and for random effects of rows and columns nested within complete replicates. Using the same definitions as before, we may write the mixed model as

$$y = X_1\beta_1 + X_2\beta_2 + Zu + e, \quad (20)$$

where the fixed effects of interest are β_1 , while we want to correct for fixed effects β_2 and random effects u . Thus, we compute corrected data

$$y_c = y - X_2\hat{\beta}_2 - Z\hat{u}, \quad (21)$$

where $\hat{\beta}_2$ and \hat{u} denote the BLUE of β_2 and the BLUP of u , respectively, based on (20). We then fit the model

$$y_c = X_1\beta_1 + e_c \quad (22)$$

by ordinary least squares.

Example 2: Table 3 reports data from a field experiment to test 35 wheat genotypes using a resolvable row-column design. A complete replicate is subdivided into incomplete rows and columns. Thus, we have incomplete blocks in both rows and columns. We may either take rows and columns as fixed (intra-block analysis) or random (combined intra-block-inter-block analysis with recovery of information). We here chose the latter option.

Table 3: Design and grain yield data for 35 genotypes of wheat in two replicates with five rows and seven columns (reproduced from Table 4.16 of Mead (1997, p. 62). § Numbers refer to genotype IDs.

	Design by Genotype ID							Yield						
	Col-1	Col-2	Col-3	Col-4	Col-5	Col-6	Col-7	Col-1	Col-2	Col-3	Col-4	Col-5	Col-6	Col-7
Replicate 1														
Row 1	20	4	33	28	7	12	30	3.77	3.21	4.55	4.09	5.05	4.19	3.27
Row 2	10	14	16	21	31	6	18	3.44	4.30	.	3.86	3.26	4.30	3.72
Row 3	22	11	19	26	29	15	23	3.49	4.20	4.77	2.56	2.87	1.93	2.26
Row 4	24	25	5	32	2	27	8	3.62	4.52	4.23	3.76	3.61	3.62	4.01
Row 5	17	9	3	34	13	35	1	3.81	3.75	4.81	3.69	4.61	2.68	4.15
Replicate 2														
Row 1	31	19	25	34	20	8	6	4.70	7.37	5.03	5.33	5.73	4.70	5.63
Row 2	24	21	12	4	23	13	3	4.07	5.66	4.98	4.04	4.27	4.10	4.75
Row 3	11	7	26	5	35	10	30	5.66	6.43	4.59	5.20	4.83	4.70	4.23
Row 4	33	9	17	18	32	15	2	5.71	6.13	4.63	5.48	5.47	.	4.16
Row 5	1	27	16	29	14	28	22	5.22	6.16	4.20	4.66	5.54	3.81	3.60

The mixed model used for analysis is

$$y_{ijk} = \mu_i + \gamma_j + r_{jh} + c_{jk} + e_{ijk} \quad (23)$$

where

y_{ijk} = yield of i -th genotype in h -th row and k -th column nested within j -th complete replicate

μ_i = mean of i -th treatment

γ_j = effect of j -th complete replicate, subject to the constraint $\sum_j \gamma_j = 0$

r_{jh} = effect of h -th row within j -th replicate; $r_{jh} \sim N(0, \sigma_r^2)$

c_{jk} = effect of k -th column within j -th replicate; $c_{jk} \sim N(0, \sigma_c^2)$

e_{ijk} = residual plot error associated with y_{ijk} ; $e_{ijk} \sim N(0, \sigma_e^2)$

The REML estimates of the variance components are $\hat{\sigma}_r^2 = 0.064$, $\hat{\sigma}_c^2 = 0.192$, and $\hat{\sigma}_e^2 = 0.090$. The adjusted genotype means obtained by a REML-based analysis using the model (20) are shown in Table 4. The same estimates are obtained by our step-wise approach as follows:

(i) We first fit (23) to obtain the BLUPs of row and column effects (Table 5) and the BLUEs of γ_j ($\hat{\gamma}_1 = -\hat{\gamma}_2 = -0.611$).

(ii) Using these, we compute corrected data $y_{c(ijkh)} = y_{ijhk} - \hat{\gamma}_j - \hat{r}_{jh} - \hat{c}_{jk}$ (Table 6).

(iii) Computing arithmetic means per genotype for these corrected data, corresponding to an ordinary least squares fit of the model $y_{c(ijkh)} = \mu_i + e_{c(ijkh)}$, yields the same means as the adjusted means in Table 4.

Table 4: Adjusted genotype means for wheat data in Table 3 based on model (20).

ID	Mean	ID	Mean	ID	Mean	ID	Mean	ID	Mean
1	4.814	8	4.602	15	3.216	22	4.012	29	3.794
2	3.915	9	4.351	16	3.955	23	3.424	30	3.953
3	5.098	10	4.328	17	4.153	24	3.888	31	3.859
4	3.521	11	4.932	18	4.564	25	4.640	32	4.264
5	4.396	12	4.946	19	5.670	26	3.759	33	4.914
6	5.410	13	4.682	20	4.319	27	4.670	34	4.298
7	5.085	14	4.765	21	4.594	28	4.294	35	3.602

Table 5: BLUPs of row and column effects (r_{jh} and c_{jk}) based on model (23) using REML.

No	Replicate 1		Replicate 2	
	Row	Column	Row	Column
1	0.164	0.089	0.168	-0.019
2	-0.113	0.151	-0.305	0.786
3	-0.291	0.137	0.156	-0.197
4	0.148	-0.035	0.149	0.198
5	0.093	0.205	-0.169	0.441
6		-0.384		-0.647
7		-0.163		-0.562

Example 2 (continued): Now assume that the genotype effect is modeled as random and we want to estimate genotype means. Thus, in model (23), we replace μ_i with $\mu + g_i$, where μ is a general mean (fixed) and g_i is the random effect of the i -th genotype (breeding values), which can be modeled to be correlated, either using pedigree information, or using marker data as in genomic selection (Piepho, 2009). In this case, we can make the following associations with respect to the general model (2): $\beta_1 =$ fixed intercept, $\beta_2 =$ fixed replicate effect, $u_1 =$ random genotype effects, $u_2 =$ random block effects. Corrected data are computed by subtracting estimates of replicate and block effects. In stage two, ordinary least squares is not the preferred method of estimation, because $V_c \neq \sigma_e^2 I$. Instead, we now need to use an estimate of V_c obtained from the first stage in order to estimate both β_1 and u_2 as per equations (11) and (12).

Table 6: Corrected data $y_{c(ijkh)} = y_{ijhk} - \hat{\gamma}_j - \hat{r}_{jh} - \hat{c}_{jk}$ for wheat data in Table 3. Fixed effects (γ_j) were estimated by BLUE (subject to the constraint $\sum_{j=1}^J \gamma_j = 0$), random effects (r_{jh}, c_{jk}) were estimated by BLUP (Table 5).

	Design by Genotype ID							Adjusted Yield						
	Col-1	Col-2	Col-3	Col-4	Col-5	Col-6	Col-7	Replicate 1						
	Col-1	Col-2	Col-3	Col-4	Col-5	Col-6	Col-7	Col-1	Col-2	Col-3	Col-4	Col-5	Col-6	Col-7
Row 1	20	4	33	28	7	12	30	4.129	3.506	4.860	4.573	5.293	5.021	3.881
Row 2	10	14	16	21	31	6	18	4.075	4.873	.	4.619	3.779	5.408	4.607
Row 3	22	11	19	26	29	15	23	4.304	4.951	5.535	3.497	3.567	3.216	3.325
Row 4	24	25	5	32	2	27	8	3.994	4.832	4.556	4.258	3.868	4.467	4.636
Row 5	17	9	3	34	13	35	1	4.240	4.117	5.191	4.244	4.923	3.582	4.831
								Replicate 2						
Row 1	31	19	25	34	20	8	6	3.940	5.805	4.448	4.353	4.510	4.567	5.413
Row 2	24	21	12	4	23	13	3	3.782	4.568	4.871	3.536	3.522	4.440	5.005
Row 3	11	7	26	5	35	10	30	4.912	4.878	4.020	4.235	3.622	4.580	4.025
Row 4	33	9	17	18	32	15	2	4.968	4.584	4.067	4.522	4.269	.	3.962
Row 5	1	27	16	29	14	28	22	4.796	4.932	3.955	4.020	4.657	4.014	3.720

DISCUSSION

This paper has shown how mixed model analyses involving fixed and random effects of interest as well as fixed and random nuisance effects that need to be corrected for can be represented in a two-stage form, where an adjustment for the nuisance effects takes place in the first stage, while estimators of the effects of interest are computed in the second stage. This two-stage representation provides a very intuitive insight into the types of adjustment involved in mixed model analysis. We found this representation to be particularly useful for illustrating adjustments in blocked experiments. In Example 1, the adjustments of the raw data when using a mixed model with random blocks are smaller than when fixed block effects are fitted because the BLUPs of block effects b_j ($j = 1, 2, 3, 4$) are shrunk towards zero (Table 1). This nicely reflects the recovery of inter-block information effected by taking blocks as random. To further illustrate this key point, it is worth considering the two extreme cases with respect to the value taken by variance component for blocks (σ_b^2) relative to that for error (σ_e^2). For balanced incomplete block designs, the BLUE of a treatment contrast $\tau = \sum_i c_i \mu_i$ ($\sum_i c_i = 0$) can be written (Yates, 1940; Mead, 1997)

$$\hat{\tau}_{comb} = \left(\frac{E \hat{\tau}_{intra}}{\sigma_e^2} + \frac{(1-E) \hat{\tau}_{inter}}{\sigma_e^2 + k \sigma_b^2} \right) / \left(\frac{E}{\sigma_e^2} + \frac{1-E}{\sigma_e^2 + k \sigma_b^2} \right), \quad (24)$$

where E is the efficiency factor ($0 < E < 1$), k is the block size, and $\hat{\tau}_{intra}$ and $\hat{\tau}_{inter}$ are the intra-block and inter-block estimators of the treatment contrast. If the block variance is zero ($\sigma_b^2 = 0$), then there is no need for any block adjustments and the BLUE of treatment means are equivalent to simple arithmetic means of raw data y_{ij} for a treatment, corresponding to an analysis based on the model without block effect ($y_{ij} = \mu_i + e_{ij}$). In this case, both types of information receive weights depending only on the average efficiency factor ($\hat{\tau}_{comb} = E\hat{\tau}_{intra} + (1-E)\hat{\tau}_{inter}$). Conversely, when the block variance becomes very large relative to the error variance ($\sigma_b^2/\sigma_e^2 \rightarrow \infty$), and hence the block adjustment is maximal, the block effect essentially becomes a fixed effect, because the mixed model equations for blocks then coincide with the ordinary least squares equations for the fixed effects model (Searle et al., 1992, p.276), meaning that all information is intra-block information ($\hat{\tau}_{comb} = \hat{\tau}_{intra}$). In practice, when error and block variance are of the same order of magnitude, the block adjustment is intermediate, reflecting the optimally weighted combination of inter-block and intra-block information. Our proposed method shows that this intermediate nature of the combined analysis corresponds to shrinkage of the estimates of block effect estimates. While this same insight is hidden in classical equations such as (24), we think that it becomes much more transparent by inspecting our equation for the corrected data.

In this paper we have assumed in all examples that $\text{var}(e) = \sigma_e^2 I$, which leads to ordinary least squares estimates in stage two of our approach in special cases. In some applications, such as analysis of repeated measures or spatial analysis of field trials, the errors associated with the observational unit have non-zero correlations and so deviate from this simple structure. If the residual error has two components, i.e., one that has a spatial or temporal correlation structure, and one that represents 'white noise' or 'nugget' (Piepho et al., 2008), then the latter component can be associated with the error term such that $\text{var}(e) = \sigma_e^2 I$, while the correlated component is assigned to the random (nuisance) effects u_2 or u . With these assignments, the same simplified results can be obtained as in the examples considered in this paper.

There is an important difference between our approach and the usual method of sweeping (De Hoog et al., 1990). For example, in the fixed-effects linear model (13), the method of sweeping computes "corrected data" $y_d = M_2 y$ with $M_2 = I - X_2 (X_2^T X_2)^{-1} X_2^T$ to take out the fixed effect β_2 . Then the reduced model is $E(y_d) = M_2 X_1 \beta_1$ and $\text{var}(y_d) = M_2 \sigma^2$, from which the reduced normal equations for β_1 are $(X_1^T M_2) M_2^- (M_2 X_1 \beta_1) = (X_1^T M_2) M_2^- (M_2 y)$, where M_2^- denotes a generalized inverse of M_2 , which simplifies to $X_1^T M_2 X_1 \beta_1 = X_1^T M_2 y$ (Searle, 1987, p.263). While this also yields the least squares estimator of β_1 under the model (13), the corrected data y_d have different expectation than the corrected data y_c in (14), for which $E(y_c) = X_1 \beta_1$. The difference arises because in our approach we use the full model, and hence both X_1 and X_2 to obtain the least squares estimate of β_2 , which is then used to correct the data, whereas the method of sweeping only uses X_2 to sweep out β_2 . While the method of sweeping has several advantages compared to ours, we find it more difficult to use in teaching non-statisticians, mainly because the reduced model for y_d looks more complicated than that for y_c . Also, our method allows corrected data to be handled as if nuisance effects had never been there in

the first place. By contrast, corrected data obtained by sweeping require accounting for the correction.

ACKNOWLEDGEMENTS

The German Federal Ministry of Education and Research (BMBF) funded this research within the AgroClustEr “Synbreed – Synergistic plant and animal breeding” (Grant ID: 0315526). We thank two anonymous referees for helpful comments on the paper.

REFERENCES

- De Hoog, F.R., Speed, T.P., Williams, E.R. (1990). On a matrix identity associated with generalized least squares. *Linear Algebra and its Applications* 127, 449–456.
- McLean, R.A., Sanders, W.L., Stroup, W.W. (1991). A unified approach to mixed linear models. *The American Statistician* 45, 54–64.
- Mead, R. (1997). Design of plant breeding trials. In: Kempton, R.A., Fox, P.N. (Eds.), *Statistical methods for plant variety evaluation*, Chapman & Hall, London, 40–67.
- Mead, R., Curnow, R.N., Hasted, A.M. (1993). *Statistical methods in agriculture and experimental biology* (2nd edition), Chapman & Hall, London.
- Patterson, H.D., Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika* 58, 545–554.
- Piepho, H.P., Richter, C., Williams, E.R. (2008). Nearest neighbour adjustment and linear variance models in plant breeding trials. *Biometrical Journal* 50, 164–189.
- Piepho, H.P. (2009). Ridge regression and extensions for genome-wide selection in maize. *Crop Science* 49, 1165–1176.
- Rao, C.R., Toutenburg, H., Shalabh, Heumann, C. (2008). *Linear models and generalizations. Least squares and alternatives* (3rd extended ed.), Springer, Berlin.
- Searle, S.R. (1987). *Linear models for unbalanced data*, Wiley, New York.
- Searle, S.R., Casella, G., McCulloch, C.E. (1992). *Variance components*, Wiley, New York.
- Yates, F. (1940). The recovery of inter-block information in balanced incomplete block designs. *Annals of Eugenics* 10, 317–325.

APPENDIX

Proof of Proposition 1:

$$\begin{aligned}
y_c &= y - X_2 \hat{\beta}_2 - Z_2 \hat{u}_2 \\
&= y - X_2 (C_{21} X_1^T V^{-1} y + C_{22} X_2^T V^{-1} y) - Z_2 G_2 Z_2^T V^{-1} (y - X \hat{\beta}) \\
&= [I - (X_2 C_{21} X_1^T V^{-1} + X_2 C_{22} X_2^T V^{-1}) - (V - V_c) P] y \\
&= [VP + (X_1 C_{11} X_1^T V^{-1} + X_1 C_{12} X_2^T V^{-1}) - VP + V_c P] y \\
&= X_1 \hat{\beta}_1 + V_c P y
\end{aligned}$$

where

$$\begin{aligned}
P &= V^{-1} - V^{-1} X (X^T V^{-1} X)^{-1} X^T V^{-1} \\
&= V^{-1} - V^{-1} (X_1 C_{11} X_1^T + X_1 C_{12} X_2^T + X_2 C_{21} X_1^T + X_2 C_{22} X_2^T) V^{-1} .
\end{aligned}$$

With this result, we find

$$\begin{aligned}
X_1 \tilde{\beta}_1 &= X_1 (X_1^T V_c^{-1} X_1)^{-1} X_1^T V_c^{-1} y_c \\
&= X_1 (X_1^T V_c^{-1} X_1)^{-1} X_1^T V_c^{-1} (X_1 \hat{\beta}_1 + V_c P y) \\
&= X_1 (X_1^T V_c^{-1} X_1)^{-1} X_1^T V_c^{-1} X_1 \hat{\beta}_1 = X_1 \hat{\beta}_1,
\end{aligned}$$

where we have used $X_1^T P = 0$ and $X_1 (X_1^T V_c^{-1} X_1)^{-1} X_1^T V_c^{-1} X_1 = X_1$ (c.f. Rao et al., 2008,

p.218). Thus, $\tilde{\beta}_1$ and $\hat{\beta}_1$ are identical. Moreover,

$$\begin{aligned}
\tilde{u}_1 &= G_1 Z_1^T V_c^{-1} (y_c - X_1 \tilde{\beta}_1) \\
&= G_1 Z_1^T V_c^{-1} (X_1 \hat{\beta}_1 + V_c P y - X_1 \hat{\beta}_1) \\
&= G_1 Z_1^T P y \\
&= G_1 Z_1^T V^{-1} (y - X \hat{\beta}) \\
&= \hat{u}_1 .
\end{aligned}$$