Review of the REBLUP method for estimating variance components under the nested error model*

Betsabé Pérez Garrido
(corresponding author)
Assistant Professor
Corvinus University
of Budapest,
Hungary
E-mail: perez.betsabe@uni-corvinus.hu

Szabolcs Szilárd Sebrek
Assistant Professor
Corvinus University
of Budapest,
Hungary
E-mail: sebrek@uni-corvinus.hu

The present work aims to analyse the REBLUP (robust empirical best linear unbiased prediction) method as proposed by *Sinha–Rao* [2009] for computing robust estimators of variance components under the nested error unit-level model. It explains the theoretical and computational aspects associated with the REBLUP method to reveal the strengths and weaknesses of the proposed approach. A Monte Carlo study is then conducted to analyse the method's performance under different scenarios.

KEYWORDS:
Linear mixed model.
Random effects.
Variance estimation.

DOI: 10.35618/hsr2019.01.en090

* The authors appreciate the helpful comments from *Isabel Molina*, *Daniel Peña*, Editor-in-Chief *Tamás Dusek* and the anonymous reviewers. *Betsabé Pérez Garrido* would like to thank the Corvinus Institute for Advanced Studies, Corvinus University of Budapest for the research fellow grant. *Szabolcs Szilárd Sebrek* gratefully acknowledges the financial support from the Hungarian National Research, Development and Innovation Office (PD16-121037).

LMMs (linear mixed models) extend simple linear models to allow for both fixed and random effects. Specifically, if an LMM contains only one random effect, the resulting model is a nested error model. Suitable data for a mixed-model analysis can be organised into different levels or clusters, also called 'multilevel' or 'hierarchical' structures. Under this framework, observations within a cluster or group are typically assumed as dependent, while the clusters themselves are assumed as independent from one another (*Gurka–Lloyd* [2007]).

Several methods exist for fitting LMMs, such as the ML (maximum likelihood), REML (restricted maximum likelihood), or moment methods (e.g. *McCulloch–Searle* [2001], *Jiang* [1996]). However, these approaches can be seriously affected by the presence of unusual observations in the data or 'outliers' – that might appear as measurement errors, random effects, or both (*Fellner* [1986], *Stahel–Welsh* [1997], *Sinha–Rao* [2009]). A small branch of literature has also focused on the problem with the robust estimation of variance components (e.g. *Pérez et al.* [2017]; *Pérez* [2011]; *Pérez–Peña–Molina* [2011]; *Molina–Peña–Pérez* [2009]; *Sinha–Rao* [2009]; *Gervini–Yohai* [1998]; *Richardson–Welsh* [1997]; *Rocke* [1983], [1991]; *Fellner* [1986]). Among these approaches, probably the most popular is the REBLUP method as proposed by *Sinha–Rao* [2009].

The present work aims to analyse the theoretical and computational aspects associated with this proposed method (*Sinha–Rao* [2009]) to demonstrate its advantages and disadvantages.

This work is organised as follows: Section 1 presents the nested error model. Section 2 introduces the EBLUP (empirical best linear unbiased prediction) estimators. Section 3 describes the problem of outliers in the data and introduces the REBLUP method. Section 4 conducts a Monte Carlo study to analyse the REBLUP method's performance under different scenarios. Section 5 demonstrates an application of this method using agricultural data; Section 6 concludes.

1. The model

Consider that our data comes from D different population groups, and suppose that there are n_d observations from group d, $d = 1, \ldots, D$, where $n = \sum_{d=1}^{D} n_d$ is the total sample size. We denote y_{dj} as the value of the study variable for the *j*-th sample unit from the *d*-th group, and $\mathbf{x}_{dj}$ is a (column) vector containing the values of *p* auxiliary variables for the same unit. The nested error model is defined as

$$y_{dj} = \mathbf{x}_{dj}^T \boldsymbol{\beta} + u_d + e_{dj}, \qquad j = 1, \ldots, n_d, \qquad d = 1, \ldots, D, \qquad /1/$$

where $\boldsymbol{\beta}$ is the $p \times 1$ vector of fixed parameters, u_d is the random effect of the *d*-th group and e_{dj} is the model error. Random group effects and errors are assumed as independent, with distributions

$$u_d \overset{iid}{\sim} N\left(0, \sigma_u^2\right) \quad \text{and} \quad e_{dj} \overset{iid}{\sim} N\left(0, \sigma_e^2\right).$$

In matricial form, let us define vector $\mathbf{y} = \left(y_{11}, y_{12}, \ldots, y_{Dn_D}\right)^T$ of size *n*; vector $\mathbf{u} = \left(u_1, u_2, \ldots, u_D\right)^T$ of size *D;* and vector $\mathbf{e} = \left(e_{11}, e_{12}, \ldots, e_{Dn_D}\right)^T$ of size *n*. The predictor vectors are given by an $n \times p$ matrix, $\mathbf{X} = \left(\mathbf{x}_{11}, \mathbf{x}_{12}, \ldots, \mathbf{x}_{Dn_D}\right)^T$, and we define an $n \times D$ block diagonal matrix as

$$\mathbf{Z} = \begin{pmatrix} 1_{n_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1_{n_D} \end{pmatrix},$$

where 1_{n_i} denotes a vector of ones of size n_i.

Thus, the model in matricial form can be written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}, \text{ where } \mathbf{u} \sim N\left(0, \sigma_u^2 \mathbf{I}_D\right), \quad \mathbf{e} \sim N\left(0, \sigma_e^2 \mathbf{I}_n\right). \qquad /2/$$

The expectation and covariance matrix of **y** are given by

$$E\left(\mathbf{y}\right) = \mathbf{X}\boldsymbol{\beta} \text{ and } var\left(\mathbf{y}\right) = \sigma_u^2 \mathbf{Z}\mathbf{Z}^T + \sigma_e^2 \mathbf{I}_n.$$

The distribution of the response variable is:

$$\mathbf{y} \sim N\left(\mathbf{X}\boldsymbol{\beta}, \sigma_u^2 \mathbf{Z}\mathbf{Z}^T + \sigma_e^2 \mathbf{I}_n\right).$$

One of the most noteworthy advantages of mixed models is that they are appropriate for non-independent data. As an illustration, consider the unit-level Model /1/ that considers data from D groups, in which these groups are independent from one another. For example, a group can be a person, a family, a county, etc. When multiple observations are collected from the same group, such as a person, a family, or a county, independence among observations from the same group can no longer be assumed. Therefore, Model /1/ adds an additional source of variation as represented by random effects u_d, to consider the data's particular structure (*Gurka–Lloyd* [2007]).

2. EBLUP estimators

The LMM defined in /2/ contains three parameters: the vector of fixed effects, $\boldsymbol{\beta}$; the vector of random group effects, **u**; and the vector of variance components, $\boldsymbol{\theta} = \left(\sigma_u^2, \sigma_e^2\right)^T$.

Assuming that the vector of variance components, $\boldsymbol{\theta}$ is known, *Henderson* [1975] notes that we can obtain BLUE (best linear unbiased estimator) of $\boldsymbol{\beta}$ and BLUP (best linear unbiased predictor) of **u**, defined respectively as

$$\tilde{\boldsymbol{\beta}} = \left(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{y} \text{ and } \tilde{\mathbf{u}} = \sigma_u^2\mathbf{Z}^T\mathbf{V}^{-1}\left(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}\right). \qquad /3/$$

Note that the estimators in /3/ depend on the vector of variance components, $\boldsymbol{\theta}$ (through matrix **V**). In practice, the vector of variance components is unknown and must be estimated from the sample data. Thus, the empirical versions of /3/ – called EBLUE (empirical best linear unbiased estimator) and EBLUP – are obtained by replacing a suitable estimator $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$, or

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}^T\hat{\mathbf{V}}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\hat{\mathbf{V}}^{-1}\mathbf{y} \text{ and } \hat{\mathbf{u}} = \hat{\sigma}_u^2\mathbf{Z}^T\hat{\mathbf{V}}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}), \qquad /4/$$

where $\hat{\mathbf{V}}$ indicates that $\boldsymbol{\theta}$ has been replaced by its estimator $\hat{\boldsymbol{\theta}}$.

Classical methods for estimating $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}$ include the ML, REML, or by-moment methods (e.g. *McCulloch–Searle* [2001], *Jiang* [1996]). In the following subsection, we present the estimation of $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}$ via ML.

2.1. Estimation via maximum likelihood

Under the ML approach and assuming the normality of **u** and **e** (*McCulloch–Searle* [2001] p. 179.) we can write the joint probability density function of **y** as

$$f\left(\boldsymbol{\beta}, \boldsymbol{\theta} \middle| \mathbf{y}\right) = (2\pi)^{-\frac{n}{2}} \left|\mathbf{V}\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\right)^T \mathbf{V}^{-1}\left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\right)\right\},$$

where the joint log-likelihood is:

$$\ell\left(\boldsymbol{\beta}, \boldsymbol{\theta} \,|\, \mathbf{y}\right) = \ln\left(f\left(\boldsymbol{\beta}, \boldsymbol{\theta} \middle| \mathbf{y}\right)\right) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln\left|\mathbf{V}\right| - \frac{1}{2}\left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\right)^T \mathbf{V}^{-1}\left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\right).$$

The first derivatives of ℓ with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ are given by

$$\frac{\delta \ell(\boldsymbol{\beta}, \boldsymbol{\theta} \,|\, \mathbf{y})}{\delta \boldsymbol{\beta}} = \mathbf{X}^T \mathbf{V}^{-1}\left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\right),$$

$$\frac{\delta \ell(\boldsymbol{\beta}, \boldsymbol{\theta} \,|\, \mathbf{y})}{\delta \boldsymbol{\theta}} = -\frac{1}{2}\frac{\delta\left(\ln\left|\mathbf{V}\right|\right)}{\delta \boldsymbol{\theta}} - \frac{1}{2}\left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\right)^T \frac{\delta \mathbf{V}^{-1}}{\delta \boldsymbol{\theta}}\left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\right),$$

which equates to zero and uses properties /1/ and /2/ as noted in the Appendix; thus, the ML equations for $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ are given as

$$\mathbf{X}^T \mathbf{V}^{-1}\left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\right) = 0, \qquad /5/$$

$$\frac{1}{2}\left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\right)^T \mathbf{V}^{-1}\frac{\delta \mathbf{V}}{\delta \boldsymbol{\theta}}\mathbf{V}^{-1}\left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\right) - \frac{1}{2}\operatorname{trace}\left\{\mathbf{V}^{-1}\frac{\delta \mathbf{V}}{\delta \boldsymbol{\theta}}\right\} = 0.$$

The equations in /5/ do not have direct solutions and must be solved numerically. Literature has provided some useful algorithms for computing the ML estimators of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$, such as the Fisher-Scoring or Newton–Raphson methods.

3. REBLUP estimators

One of the most important disadvantages of the ML equations in /5/ is that they are sensitive to the presence of outliers in the data (*Fellner* [1986], *Stahel–Welsh* [1997], *Sinha–Rao* [2009]). To overcome this disadvantage, *Sinha* and *Rao* [2009] proposed a more resistant version of the ML equations against outlier observations, called the REBLUP method; essentially, if some fitted values unusually differ from the corresponding observed values, then this indicates apparent outliers in the data. Therefore, *Sinha* and *Rao* [2009] proposed a robust version of the ML equations to handle outliers in the response values, given by the following expressions:

$$\mathbf{X}^T\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi(\mathbf{r}) = 0,$$

$$\psi^T(\mathbf{r})\mathbf{U}^{\frac{1}{2}}\mathbf{V}^{-1}\frac{\delta\mathbf{V}}{\delta\boldsymbol{\theta}}\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi(\mathbf{r}) - \text{trace}\left\{\mathbf{K}\mathbf{V}^{-1}\frac{\delta\mathbf{V}}{\delta\boldsymbol{\theta}}\right\} = 0, \qquad /6/$$

where $\mathbf{r}$ denotes the standardised residuals $\mathbf{r} = \mathbf{U}^{-\frac{1}{2}}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$; $\mathbf{U} = \text{diag}(\mathbf{V})$; and $\mathbf{K} = c\mathbf{I}_n$ with $c = E\left[\psi_b^2(\mathbf{r})\right]$. In the former expressions, $\psi(\cdot)$ represents a smoothing function (e.g. Huber's psi function, Tukey's biweight, or Hampel's function) and is used for smoothing all observations with large residuals, as these indicate apparent outliers in the data. Particularly, the REBLUP method considers *Huber*'s [1964] psi function, defined as $\psi_b(\text{u}) = u \cdot \min\left(1, \frac{b}{|u|}\right)$ with the turning constant $b = 1.345$ to reach 95% efficiency.

The REBLUP estimators are obtained based on a two-step procedure that uses the Newton–Raphson algorithm and takes ML estimators as starting values, as follows:

Stage 1. Estimate $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ simultaneously based on the robust ML equations in /6/.

Stage 2. Predictor $\mathbf{u}$ is obtained using the estimator in Stage 1.

Stage 1 simultaneously estimates two model parameters ($\boldsymbol{\beta}$ and $\boldsymbol{\theta}$). Hence, to be able to analyse the estimated vector of variance components, $\boldsymbol{\theta}$, we also need to introduce the estimation of the vector of fixed effects, $\boldsymbol{\beta}$.

3.1. Details of the method and our contribution

Sinha and *Rao*'s [2009] study present the ML equations given in /5/, the robust version of the ML equations given in /6/, and the iterative equations given in /7/ and /8/. Our work aims to give more details about the construction of the REBLUP method (i.e. introduces the steps [for instance, the derivatives of expressions /7/

and /8/] omitted in *Sinha* and *Rao*'s [2009] study). This information could be useful for researchers, programmers and final users.

3.2. Stage 1

During the first stage, two parameters of Model /2/, $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ are estimated simultaneously. In the next subsection, we present the estimation of the vector of fixed effects, $\boldsymbol{\beta}$.

3.2.1. Estimation of the fixed effects, β

Consider the first robust ML equation in /6/ and denote $S\left(\hat{\boldsymbol{\beta}}\right) = \mathbf{X}^T\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi\left(\mathbf{r}\right)$. The Newton–Raphson iterative equation for the estimation of $\boldsymbol{\beta}$ is given by the expression

$$\hat{\boldsymbol{\beta}} \approx \boldsymbol{\beta} - \left[\frac{\delta S\left(\hat{\boldsymbol{\beta}}\right)}{\delta \hat{\boldsymbol{\beta}}} |_{\hat{\boldsymbol{\beta}} = \boldsymbol{\beta}} \right]^{-1} S\left(\boldsymbol{\beta}\right),$$

where the derivative is:

$$\frac{\delta S\left(\hat{\boldsymbol{\beta}}\right)}{\delta \hat{\boldsymbol{\beta}}} |_{\hat{\boldsymbol{\beta}} = \boldsymbol{\beta}} = \frac{\delta \left(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi\left(\mathbf{r}\right) \right)}{\delta \hat{\boldsymbol{\beta}}} |_{\hat{\boldsymbol{\beta}} = \boldsymbol{\beta}} = \\ = -\mathbf{X}^T\mathbf{V}^{-1} \operatorname{diag}\left(\frac{\delta\psi\left(r_1\right)}{\delta \hat{\boldsymbol{\beta}}}, \frac{\delta\psi\left(r_2\right)}{\delta \hat{\boldsymbol{\beta}}}, \ldots, \frac{\delta\psi\left(r_n\right)}{\delta \hat{\boldsymbol{\beta}}} \right)\mathbf{X} \qquad /7/$$

with $\frac{\delta\psi\left(r_i\right)}{\delta \hat{\boldsymbol{\beta}}} = \begin{cases} 1, \text{ if } \left|r_i\right| \leq b; \\ 0, \text{ otherwise.} \end{cases}$

3.2.2. Estimation of the variance components, θ

In this part of the study, the estimation of the vector of variance components, $\boldsymbol{\theta}$ is introduced. Let us consider the second robust ML equation in /6/ and denote

$$S\left(\hat{\boldsymbol{\theta}}\right) = \psi^T\left(\mathbf{r}\right)\mathbf{U}^{\frac{1}{2}}\mathbf{V}^{-1}\frac{\delta \mathbf{V}}{\delta \boldsymbol{\theta}}\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi\left(\mathbf{r}\right) - \operatorname{trace}\left\{ \mathbf{K}\mathbf{V}^{-1}\frac{\delta \mathbf{V}}{\delta \boldsymbol{\theta}} \right\}.$$

The Newton–Raphson iterative equation for the estimation of $\boldsymbol{\theta}$ is given by

$$\hat{\boldsymbol{\theta}} \approx \boldsymbol{\theta} - \left[\frac{\delta S\left(\hat{\boldsymbol{\theta}}\right)}{\delta \hat{\boldsymbol{\theta}}} \Big|_{\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}} \right]^{-1} S\left(\boldsymbol{\theta}\right). \qquad /8/$$

To calculate the derivative $\frac{\delta S\left(\hat{\boldsymbol{\theta}}\right)}{\delta \hat{\boldsymbol{\theta}}}$, let us first rewrite $S\left(\hat{\boldsymbol{\theta}}\right)$ as follows:

$$S\left(\hat{\boldsymbol{\theta}}\right) = \left(\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi\left(\mathbf{r}\right)\right)^{T} \frac{\delta \mathbf{V}}{\delta \boldsymbol{\theta}} \left(\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi\left(\mathbf{r}\right)\right) - \operatorname{trace}\left\{\mathbf{K}\mathbf{V}^{-1}\frac{\delta \mathbf{V}}{\delta \boldsymbol{\theta}}\right\}. \qquad /9/$$

The first element in /9/ is a quadratic form. Under the assumption that $\frac{\delta \mathbf{V}}{\delta \boldsymbol{\theta}}$ is a symmetric matrix, we can use properties /2/ and /3/ from the Appendix to obtain the derivative, given by

$$\frac{\delta S\left(\hat{\boldsymbol{\theta}}\right)}{\delta \hat{\boldsymbol{\theta}}} = 2\frac{\delta \mathbf{V}}{\delta \boldsymbol{\theta}} \left(\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi\left(\mathbf{r}\right)\right) \frac{\delta\left(\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi\left(\mathbf{r}\right)\right)}{\delta \boldsymbol{\theta}} + \operatorname{trace}\left\{\mathbf{K}\mathbf{V}^{-1}\frac{\delta \mathbf{V}}{\delta \boldsymbol{\theta}}\mathbf{V}^{-1}\frac{\delta \mathbf{V}}{\delta \boldsymbol{\theta}}\right\} =$$

$$= 2\frac{\delta\left(\psi\left(\mathbf{r}\right)^{T}\mathbf{U}^{\frac{1}{2}}\mathbf{V}^{-1}\right)}{\delta \boldsymbol{\theta}} \frac{\delta \mathbf{V}}{\delta \boldsymbol{\theta}} \left(\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi\left(\mathbf{r}\right)\right) + \operatorname{trace}\left\{\mathbf{K}\mathbf{V}^{-1}\frac{\delta \mathbf{V}}{\delta \boldsymbol{\theta}}\mathbf{V}^{-1}\frac{\delta \mathbf{V}}{\delta \boldsymbol{\theta}}\right\}, \qquad 10/$$

and using the product rule

$$\frac{\delta\left(\psi\left(\mathbf{r}\right)^{T}\mathbf{U}^{\frac{1}{2}}\mathbf{V}^{-1}\right)}{\delta \boldsymbol{\theta}} = \frac{\delta \psi\left(\mathbf{r}\right)^{T}}{\delta \boldsymbol{\theta}}\mathbf{U}^{\frac{1}{2}}\mathbf{V}^{-1} + \psi\left(\mathbf{r}\right)^{T}\frac{\delta \mathbf{U}^{\frac{1}{2}}}{\delta \boldsymbol{\theta}}\mathbf{V}^{-1} + \psi\left(\mathbf{r}\right)^{T}\mathbf{U}^{\frac{1}{2}}\frac{\delta \mathbf{V}^{-1}}{\delta \boldsymbol{\theta}}$$

and property /2/

$$\frac{\delta\left(\psi\left(\mathbf{r}\right)^{T}\mathbf{U}^{\frac{1}{2}}\mathbf{V}^{-1}\right)}{\delta \boldsymbol{\theta}} = \frac{\delta \psi\left(\mathbf{r}\right)^{T}}{\delta \boldsymbol{\theta}}\mathbf{U}^{\frac{1}{2}}\mathbf{V}^{-1} + \psi\left(\mathbf{r}\right)^{T}\frac{\delta \mathbf{U}^{\frac{1}{2}}}{\delta \boldsymbol{\theta}}\mathbf{V}^{-1} - \psi\left(\mathbf{r}\right)^{T}\mathbf{U}^{\frac{1}{2}}\mathbf{V}^{-1}\frac{\delta \mathbf{V}}{\delta \boldsymbol{\theta}}\mathbf{V}^{-1}. \qquad /11/$$

Expressions /9/, /10/ and /11/ are defined for both elements of the vector of variance components, namely, for the random effect associated with the groups σ_u^2, and the random effect associated with the error term σ_e^2. The following subsections will explicitly define these expressions.

Estimation of the random effect associated with the groups, σ_u^2

In the case of the random effect associated with the groups, expression /9/ is given as

$$S(\hat{\boldsymbol{\theta}}_u) = \left(\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi(\mathbf{r})\right)^T \frac{\delta \mathbf{V}}{\delta \sigma_u^2} \left(\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi(\mathbf{r})\right) - \text{trace}\left\{\mathbf{K}\mathbf{V}^{-1}\frac{\delta \mathbf{V}}{\delta \sigma_u^2}\right\},$$

considering that

$$\frac{\delta \mathbf{V}}{\delta \sigma_u^2} = \frac{\delta\left(\sigma_u^2 \mathbf{Z}\mathbf{Z}^T + \sigma_e^2 \mathbf{I}_n\right)}{\delta \sigma_u^2} = \mathbf{Z}\mathbf{Z}^T,$$

we have

$$S(\hat{\boldsymbol{\theta}}_u) = \left(\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi(\mathbf{r})\right)^T \mathbf{Z}\mathbf{Z}^T \left(\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi(\mathbf{r})\right) - \text{trace}\left\{\mathbf{K}\mathbf{V}^{-1}\mathbf{Z}\mathbf{Z}^T\right\}.$$

The derivative in /10/ is given by

$$\frac{\delta S(\hat{\boldsymbol{\theta}}_u)}{\delta \sigma_u^2} = 2\frac{\delta\left(\psi(\mathbf{r})^T \mathbf{U}^{\frac{1}{2}}\mathbf{V}^{-1}\right)}{\delta \sigma_u^2} \frac{\delta \mathbf{V}}{\delta \sigma_u^2} \left(\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi(\mathbf{r})\right) + \text{trace}\left\{\mathbf{K}\mathbf{V}^{-1}\frac{\delta \mathbf{V}}{\delta \sigma_u^2}\mathbf{V}^{-1}\frac{\delta \mathbf{V}}{\delta \sigma_u^2}\right\} =$$

$$= 2\frac{\delta\left(\psi(\mathbf{r})^T \mathbf{U}^{\frac{1}{2}}\mathbf{V}^{-1}\right)}{\delta \sigma_u^2} \mathbf{Z}\mathbf{Z}^T \left(\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi(\mathbf{r})\right) + \text{trace}\left\{\mathbf{K}\mathbf{V}^{-1}\mathbf{Z}\mathbf{Z}^T\mathbf{V}^{-1}\mathbf{Z}\mathbf{Z}^T\right\},$$

and expression /11/ is:

$$\frac{\delta\left(\psi(\mathbf{r})^T\mathbf{U}^{\frac{1}{2}}\mathbf{V}^{-1}\right)}{\delta\sigma_u^2} = \frac{\delta\psi(\mathbf{r})^T}{\delta\sigma_u^2}\mathbf{U}^{\frac{1}{2}}\mathbf{V}^{-1} + \frac{1}{2}\psi(\mathbf{r})^T\mathbf{U}^{-\frac{1}{2}}\mathbf{V}^{-1} - \psi(\mathbf{r})^T\mathbf{U}^{\frac{1}{2}}\mathbf{V}^{-1}\mathbf{Z}\mathbf{Z}^T\mathbf{V}^{-1}$$

with $\frac{\delta\psi(\mathbf{r})}{\delta\sigma_u^2} = -\frac{1}{2}\mathbf{U}^{-1}\mathbf{r}\psi(\mathbf{r})^T$.

Thus, the Newton–Raphson iterative equation for estimating σ_u^2 or

$$\hat{\sigma}_u^2 \approx \sigma_{u0}^2 - \left[\frac{\delta S(\hat{\boldsymbol{\theta}}_u)}{\delta\hat{\sigma}_u^2}|_{\hat{\sigma}_u^2=\sigma_u^2}\right]^{-1} S(\hat{\boldsymbol{\theta}}_u)$$

can be expressed as

$$\hat{\sigma}_u^2 = \sigma_{u0}^2 - \left[2\frac{\delta\left(\psi(\mathbf{r})^T\mathbf{U}^{\frac{1}{2}}\mathbf{V}^{-1}\right)}{\delta\sigma_u^2}\mathbf{Z}\mathbf{Z}^T\left(\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi(\mathbf{r})\right) + \text{trace}\left\{\mathbf{K}\mathbf{V}^{-1}\mathbf{Z}\mathbf{Z}^T\mathbf{V}^{-1}\mathbf{Z}\mathbf{Z}^T\right\}\right]^{-1} \cdot \quad /12$$

$$\cdot\left[\left(\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi(\mathbf{r})\right)^T\mathbf{Z}\mathbf{Z}^T\left(\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi(\mathbf{r})\right) - \text{trace}\left\{\mathbf{K}\mathbf{V}^{-1}\mathbf{Z}\mathbf{Z}^T\right\}\right]$$

with σ_{u0}^2 as a starting value, as *Sinha* and *Rao* [2009] note that this corresponds to an ML estimator.

Estimation of the random effect associated with the error term, σ_e^2

In the case of the random effect associated with the error term, expression /9/ is given by

$$S(\hat{\boldsymbol{\theta}}_e) = \left(\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi(\mathbf{r})\right)^T\frac{\delta\mathbf{V}}{\delta\sigma_e^2}\left(\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi(\mathbf{r})\right) - \text{trace}\left\{\mathbf{K}\mathbf{V}^{-1}\frac{\delta\mathbf{V}}{\delta\sigma_e^2}\right\},$$

considering that

$$\frac{\delta \mathbf{V}}{\delta \sigma_e^2} = \frac{\delta\left(\sigma_u^2 \mathbf{Z}\mathbf{Z}^T + \sigma_e^2 \mathbf{I}_n\right)}{\delta \sigma_e^2} = \mathbf{I}_n ,$$

we have

$$S\left(\hat{\boldsymbol{\theta}}_e\right) = \left(\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi(\mathbf{r})\right)^T \left(\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi(\mathbf{r})\right) - \text{trace}\left\{\mathbf{K}\mathbf{V}^{-1}\right\} .$$

The derivative /10/ is given by

$$\frac{\delta S\left(\hat{\boldsymbol{\theta}}_e\right)}{\delta \sigma_e^2} = 2\frac{\delta\left(\psi(\mathbf{r})^T \mathbf{U}^{\frac{1}{2}}\mathbf{V}^{-1}\right)}{\delta \sigma_e^2}\frac{\delta \mathbf{V}}{\delta \sigma_e^2}\left(\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi(\mathbf{r})\right) + \text{trace}\left\{\mathbf{K}\mathbf{V}^{-1}\frac{\delta \mathbf{V}}{\delta \sigma_e^2}\mathbf{V}^{-1}\frac{\delta \mathbf{V}}{\delta \sigma_e^2}\right\} =$$

$$= 2\frac{\delta\left(\psi(\mathbf{r})^T \mathbf{U}^{\frac{1}{2}}\mathbf{V}^{-1}\right)}{\delta \sigma_e^2}\left(\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi(\mathbf{r})\right) + \text{trace}\left\{\mathbf{K}\mathbf{V}^{-1}\mathbf{V}^{-1}\right\},$$

and expression /11/ is given by

$$\frac{\delta\left(\psi(\mathbf{r})^T \mathbf{U}^{\frac{1}{2}}\mathbf{V}^{-1}\right)}{\delta \sigma_e^2} = \frac{\delta \psi(\mathbf{r})^T}{\delta \sigma_e^2}\mathbf{U}^{\frac{1}{2}}\mathbf{V}^{-1} + \frac{1}{2}\psi(\mathbf{r})^T \mathbf{U}^{-\frac{1}{2}}\mathbf{V}^{-1} - \psi(\mathbf{r})^T \mathbf{U}^{\frac{1}{2}}\mathbf{V}^{-1}\mathbf{V}^{-1}$$

with $\frac{\delta \psi(\mathbf{r})}{\delta \sigma_e^2} = -\frac{1}{2}\mathbf{U}^{-1}\mathbf{r}\psi(\mathbf{r})^T$.

Thus, the Newton–Raphson iterative equation for estimating σ_e^2 or

$$\hat{\sigma}_e^2 \approx \sigma_{e0}^2 - \left[\frac{\delta S\left(\hat{\boldsymbol{\theta}}_e\right)}{\delta \hat{\sigma}_e^2}\Big|_{\hat{\sigma}_e^2 = \sigma_e^2}\right]^{-1} S\left(\hat{\boldsymbol{\theta}}_e\right)$$

can be expressed as

$$\hat{\sigma}_e^2 = \sigma_{e0}^2 - \left[2\frac{\delta\left(\psi(\mathbf{r})^T \mathbf{U}^{\frac{1}{2}}\mathbf{V}^{-1}\right)}{\delta\sigma_e^2}\left(\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi(\mathbf{r})\right) + \text{trace}\left\{\mathbf{K}\mathbf{V}^{-1}\mathbf{V}^{-1}\right\} \right]^{-1} \cdot$$
$$\cdot \left[\left(\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi(\mathbf{r})\right)^T \left(\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi(\mathbf{r})\right) - \text{trace}\left\{\mathbf{K}\mathbf{V}^{-1}\right\} \right] \qquad /13/$$

with σ_{e0}^2 as a starting value, as *Sinha* and *Rao* [2009] note that this also corresponds to an ML estimator.

3.3. Stage 2

In the second stage, *Sinha* and *Rao* [2009] predict **u** based on the estimators in Stage 1. However, as we aim to analyse the estimation of variance components, we will omit the details of this stage.

In conclusion, we have presented above the explicit expressions for the estimation of the vector of fixed effects, $\boldsymbol{\beta}$ given by /7/. We have defined the expressions for the estimation of the vector of variance components, $\boldsymbol{\theta}$, specifically, the estimator of the random effect associated with the groups, σ_u^2 given by /12/, and the estimator of the random effect associated with the error term, σ_e^2 given by /13/.

4. Monte Carlo simulation

This section presents a simulation study to evaluate the REBLUP method for uncontaminated data (data without outliers) and contaminated data (data with outliers). Uncontaminated data were generated from the nested error model given by

$$y_{dj} = \beta x_{dj} + u_d + e_{dj}, \qquad j = 1, ..., n_d, \qquad d = 1, ..., D,$$
$$u_d \overset{iid}{\sim} N\left(0, \sigma_u^2\right), \quad \text{and} \quad e_{dj} \overset{iid}{\sim} N\left(0, \sigma_e^2\right),$$

considering one auxiliary variable, $\beta = 1$ and the variance associated with the error term $\sigma_e^2 = 1$. We discuss four alternatives for the variance associated with the groups, or specifically, $\sigma_u^2 \in \{0.3, 0.6, 1, 1.3\}$. Our data consists of $D = 50$ groups with a constant sample size $n_d = 20$, $d = 1, \ldots, D$. Values for the auxiliary variable were generated as $x_{dj} \overset{iid}{\sim} N(0,1)$, $j = 1, \ldots, n_d$, $d = 1, \ldots, D$ and it was kept as fixed during the simulations. The scenarios include the following:

Scenario 1: Uncontaminated data (data without outliers).

Scenario 2: Contaminated data (one additive outlier of increasing size – outliers affecting the variance associated with the error term σ_e^2). This scenario considers one outlying observation of increasing size in the last area, or

$$y_{D1} = \beta x_{D1} + u_D + \left(e_{D1} + C\right), \quad C = 4k, \quad k = 0, 1, \ldots, 10.$$

Scenario 3: Contaminated data (one shifted area – outliers affecting the variance associated with the groups σ_u^2). This scenario introduces one outlying group, or specifically, we select the last group and replace all observations on that group by a mean shift of increasing size, or

$$y_{Dj} = \beta x_{Dj} + \left(u_D + C\right) + e_{Dj}, \quad j = 1, \ldots, n_D, \quad C = 4k, \quad k = 0, 1, \ldots, 10.$$

We then run our simulations separately for each of the former scenarios. In each simulation replicated, we calculate the non-robust (ML and REML) estimators and use the robust REBLUP method with ML estimators as starting values. Regarding Monte Carlo replications, $R = 200$. We denote $\hat{\theta}_{\ell,\mathrm{m}}^{(r)}$ as the resulting value of estimator $m = \{\text{ML, REML, REBLUP}\}$ of the variance component θ_ℓ, $\ell = 1, 2$, in the Monte Carlo replication *r*. For each scenario, we compute the following measures of performance:

ARB (absolute relative bias):

$$ARB\left(\hat{\theta}_{\ell,m}\right) = 100 \times \frac{\left|\left(1/R\right)\sum_{r=1}^{R} \hat{\theta}_{\ell,m}{}^{(r)} - \theta_\ell\right|}{\theta_\ell};$$

RMAE (relative mean absolute error):

$$RMAE\left(\hat{\theta}_{\ell,m}\right) = 100 \times \frac{\left(1/R\right)\sum_{r=1}^{R}\left|\hat{\theta}_{\ell,m}^{(r)} - \theta_\ell\right|}{\theta_\ell}.$$

Table 1 reports the ARB and RMAE for the considered estimators under Scenario 1 (uncontaminated data). Table 1 illustrates that the two non-robust (ML and REML) estimators are less unbiased than the robust REBLUP estimator. This result is reasonable, as the ML and REML methods were created to work well under no contamination; further, the bias of the REBLUP estimator associated with the error term σ_e^2 increases as σ_u^2 increases.

Table 1

ARB and RMAE of estimators under Scenario 1 (no contamination)

Case	Method	ARB		RMAE	
		σ_e^2	σ_u^2	σ_e^2	σ_u^2
$\sigma_u^2 = 0.3$	ML	0.11	4.29	3.50	19.34
	REML	0.21	1.99	3.52	19.36
	REBLUP	1.91	7.89	4.44	20.60
$\sigma_u^2 = 0.6$	ML	0.49	3.86	3.56	16.40
	REML	0.60	1.73	3.58	16.43
	REBLUP	2.80	7.59	4.76	17.57
$\sigma_u^2 = 1.0$	ML	0.23	5.31	3.66	17.60
	REML	0.33	3.28	3.66	17.49
	REBLUP	4.26	6.89	5.47	19.16
$\sigma_u^2 = 1.3$	ML	0.17	3.17	3.54	15.99
	REML	0.27	1.11	3.55	16.02
	REBLUP	4.93	5.59	6.17	16.78

In Scenario 2 we contaminate our data by adding one additive outlier of increasing size in an attempt to affect the estimation of the variance components associated with the error term σ_e^2. Figures A1 and A2 of the Appendix demonstrate that the ML and REML estimators produce similar values for the ARB and RMAE. Figure A1 displays two issues: *1*. the REBLUP estimator of σ_e^2 (and σ_u^2) has a larger bias than the non-robust estimators ML and REML, and *2*. this becomes more evident as we increase the outlier size. This is because this type of outlier affects the starting values of the Newton–Raphson algorithm (based on ML estimators), and consequently, the algorithm diverges. A similar conclusion is made for the RMAE of the REBLUP estimators of σ_e^2 (and σ_u^2), as noted in Figure A2.

Finally, we contaminate our data in Scenario 3 by introducing one outlying area, in which we attempt to affect the estimation of the variance components associated with the groups σ_u^2. Figures A3 and A4 again illustrate that the ML and REML estimators produce similar values for the ARB and RMAE. Figure A3 indicates that the REBLUP estimator for σ_e^2 (and σ_u^2) exhibits a larger bias than the non-robust estimators ML and REML. This type of outlier also affects the Newton–Raphson algorithm's starting values, failing to estimate both elements of the variance components (σ_u^2 and σ_u^2).

5. Application – County crop areas

Here we use the dataset presented by *Battese*, *Harter* and *Fuller* [1988], which consists of survey and satellite data for 12 counties in Iowa, in the United States (groups), or $D = 12$, with $n = 37$ number of observations. The data contain information about the number of segments in each county, number of reported hectares, number of pixels in the sample segments, and the mean number of pixels per segment, the latter of which we omitted in the present study. Table 2 displays the considered data.

Table 2

Survey and satellite data for corn and soybeans in 12 Iowa counties, 1988

Observation	County	Number of segments		Reported hectares		Number of pixels in sample segments	
		Sample	County	Corn	Soybeans	Corn	Soybeans
1	Cerro Gordo	1	545	165.76	8.09	374	55
2	Hamilton	1	566	96.32	106.03	209	218
3	Worth	1	394	76.08	103.60	253	250
4	Humboldt	2	424	185.35	6.47	432	96
5				116.43	63.82	367	178

(Continued on the next page)

(Continued)

Observation	County	Number of segments		Reported hectares		Number of pixels in sample segments	
		Sample	County	Corn	Soybeans	Corn	Soybeans
6	Franklin	3	564	162.08	43.50	361	137
7				152.04	71.43	288	206
8				161.75	42.49	369	165
9	Pocahontas	3	570	92.88	105.26	206	218
10				149.94	76.49	316	221
11				64.75	174.34	145	338
12	Winnebago	3	402	127.07	95.67	355	128
13				133.55	76.57	295	147
14				77.70	93.48	223	204
15	Wright	3	567	206.39	37.84	459	77
16				108.33	131.12	290	217
17				118.17	124.44	307	258
18	Webster	4	687	99.96	144.15	252	303
19				140.43	103.60	293	221
20				98.95	88.59	206	222
21				131.04	115.58	302	274
22	Hancock	5	569	114.12	99.15	313	190
23				100.60	124.56	246	270
24				127.88	110.88	353	172
25				116.90	109.14	271	228
26				87.41	143.66	237	297
27	Kossuth	5	965	93.48	91.05	221	167
28				121.00	132.33	369	191
29				109.91	143.14	343	249
30				122.66	104.13	342	182
31				104.21	118.57	294	179
32	Hardin	6	556	88.59	102.59	220	262
33				88.59	29.46	340	87
34				165.35	69.28	355	160
35				104.00	99.15	261	221
36				88.63	143.66	187	345
37				153.70	94.49	350	190

Source: *Battese–Harter–Fuller* [1988] p. 28.

The considered nested error model is:

$$y_{dj} = \beta_0 + \beta_1 x_{1dj} + \beta_2 x_{2dj} + u_d + e_{dj},$$

$$j = 1, \ldots, n_d, \quad d = 1, \ldots, D = 12,$$

where y_{dj} is the number of hectares of corn in the j-th segment of the d-th county; n_d is the number of sample segments in the d-th county; and x_{1dj} and x_{2dj} denote the number of pixels classified as corn and soybeans, respectively. Random group effects and errors are assumed as independent, with distributions

$$u_d \overset{iid}{\sim} N\left(0, \sigma_u^2\right) \quad \text{and} \quad e_{dj} \overset{iid}{\sim} N\left(0, \sigma_e^2\right).$$

Table 3 presents the estimates of model parameters $(\boldsymbol{\beta}, \boldsymbol{\theta})$. The vector of fixed effects is given by $\boldsymbol{\beta} = \left(\beta_0, \beta_1, \beta_2\right)$, and the vector of the variance components is $\boldsymbol{\theta} = \left(\sigma_u^2, \sigma_e^2\right)^T$. We have obtained the estimates applying two non-robust methods (ML and REML) and one robust method (REBLUP), using ML estimators as starting values.

Table 3

Estimates of model parameters $(\boldsymbol{\beta}, \boldsymbol{\theta})$

Coefficient	Estimate		
	ML	REML	REBLUP
Intercept (β_0)	18.09	17.96	29.14
Corn pixels (β_1)	0.36	0.37	0.36
Soybean pixels (β_2)	–0.03	–0.03	–0.07
Random effect associated with the error term (σ_e^2)	280.23	297.71	225.60
Random effect associated with the groups (σ_u^2)	47.80	63.31	102.74

Finally, the computational time used to calculate the estimates $(\boldsymbol{\beta}, \boldsymbol{\theta})$ presented in Table 3 is as follows: ML – 0.06 sec, REML – 0.08 sec, REBLUP – 2.73 sec. The time data for the non-robust methods (ML and REML) are less than one second, while it is around three seconds for the REBLUP method, meaning that the latter method is relatively fast.

This work resorts to the R statistical software. We created our own code (available upon request from the authors) for computing the REBLUP estimates, while package *nlme* was used to compute the ML and REML estimates.

6. Conclusion

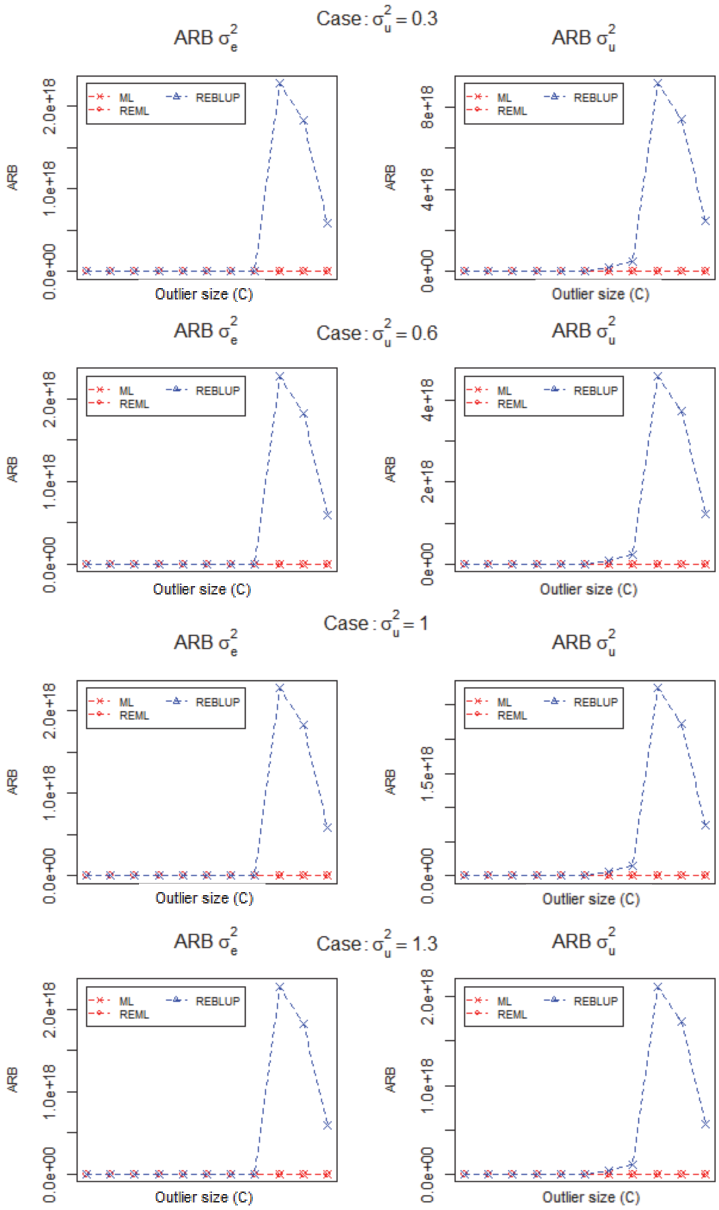
This work analysed the REBLUP method's performance in computing robust estimators of variance components under the nested error model. Our theoretical and simulation study identified some possible disadvantages that can be considered for future research.

First, during the first stage of the REBLUP method, the model parameters $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ are simultaneously estimated using the Newton–Raphson algorithm through Equations /7/, /12/, and /13/. This set of equations depends on the vector of variance components, $\boldsymbol{\theta}$ (through matrix $\mathbf{V}$), which is unknown. The REBLUP method suggests that ML estimators be used as starting values for the Newton–Raphson algorithm, then the estimators of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ are progressively and simultaneously discovered. In the presence of outliers, the starting values (based on ML estimators) can be seriously affected; consequently, the Newton–Raphson algorithm may not converge. Moreover, the type of outliers as described in Section 4 can also affect the estimation of more model parameters. As an illustration, consider Scenario 2, in which we introduced one additive outlier of increasing size to affect the variance associated with the error term σ_e^2. Figure A1 demonstrates that the bias of the REBLUP estimator associated with the error term σ_e^2 is larger. Further, the bias of the REBLUP estimator associated with the groups σ_u^2 is also larger; in other words, this type of outlier simultaneously affects two model parameters: σ_e^2 and σ_u^2. We suggest overcoming this problem by replacing the starting values (currently, the ML estimators) from the Newton–Raphson algorithm for robust starting values free of the influence of outlier observations.

Second, a different approach to improve the REBLUP method could explore a more resistant algorithm replacing the Newton–Raphson algorithm, exempt from the influence of outlier observations. Finally, we believe that future research on this topic might incorporate similar scenarios affecting directly the variances associated with the groups and errors as described in Section 4.

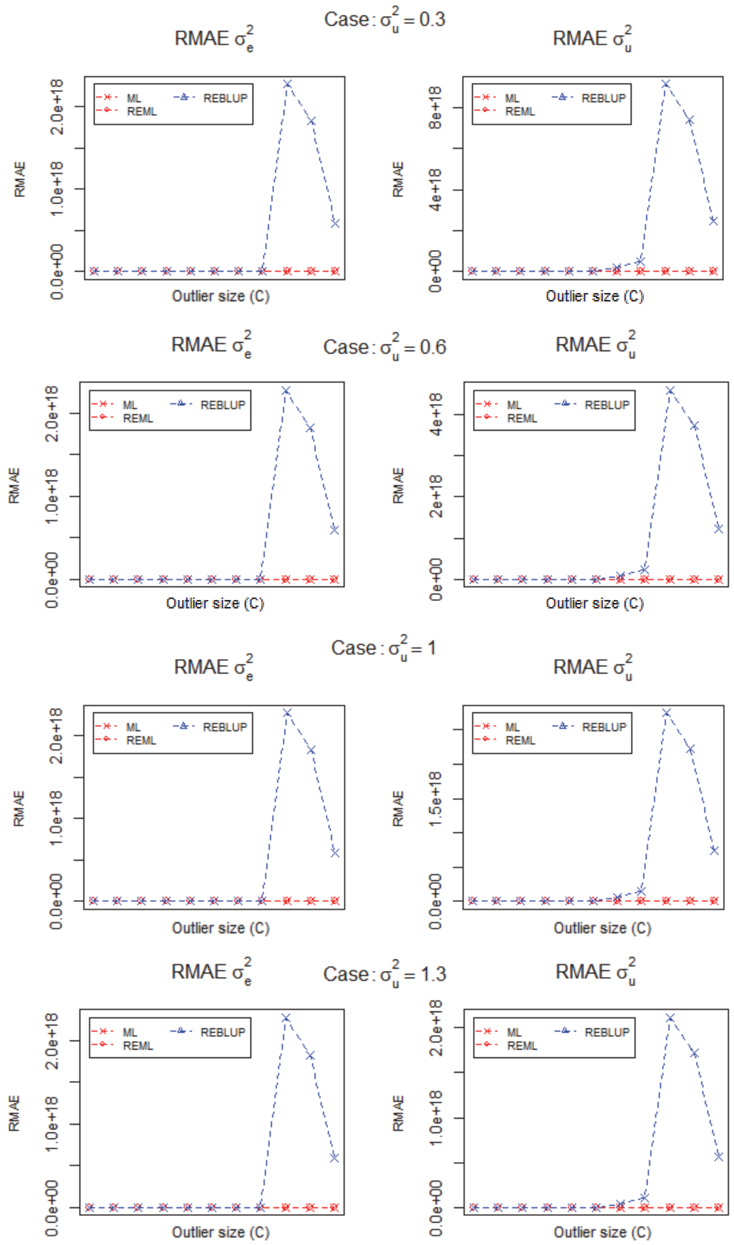
Appendix

Figure A1. ARB of ML, REML and REBLUP estimators under Scenario 2
(one additive outlier of increasing size)



Note. Here and in the following figures, ARB: absolute relative bias; ML: maximum likelihood; REML: restricted maximum likelihood; REBLUP: robust empirical best linear unbiased prediction.

Figure A2. RMAE of ML, REML and REBLUP estimators under Scenario 2 (one additive outlier of increasing size)



Note. Here and in Figure A4, RMAE: relative mean absolute error.

Figure A3. ARB of ML, REML and REBLUP estimators under Scenario 3 (one outlying area)

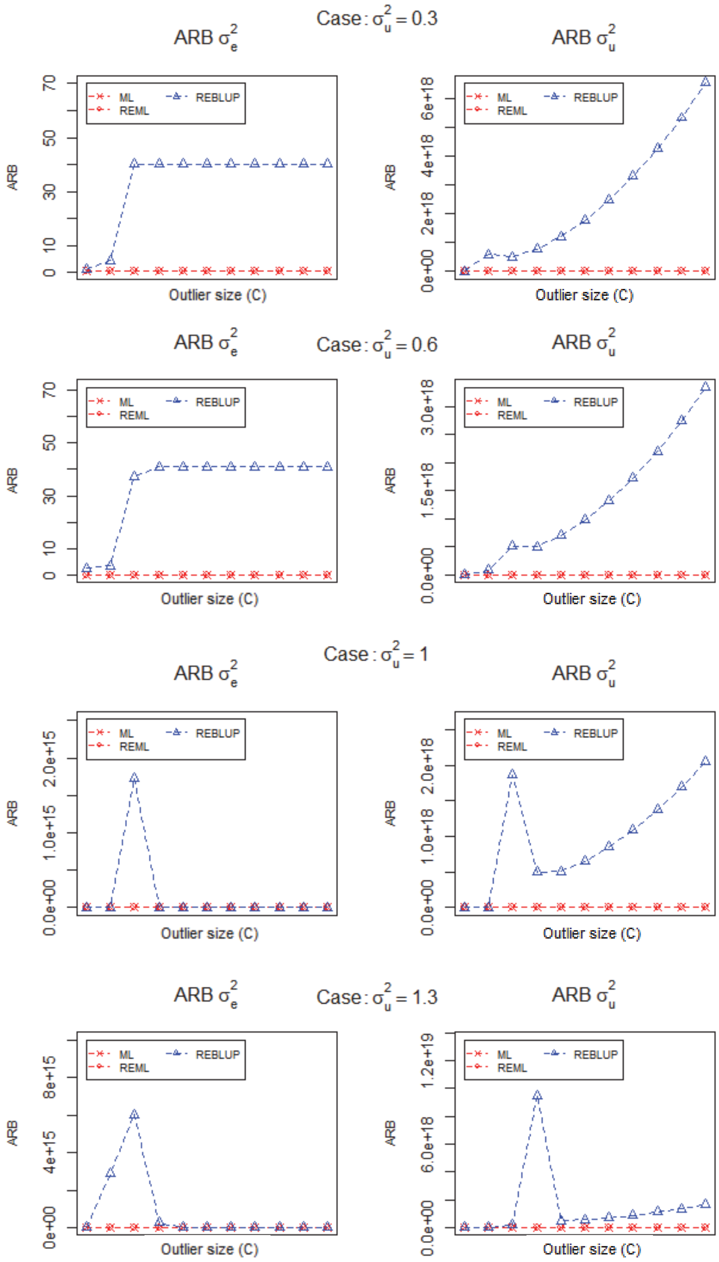
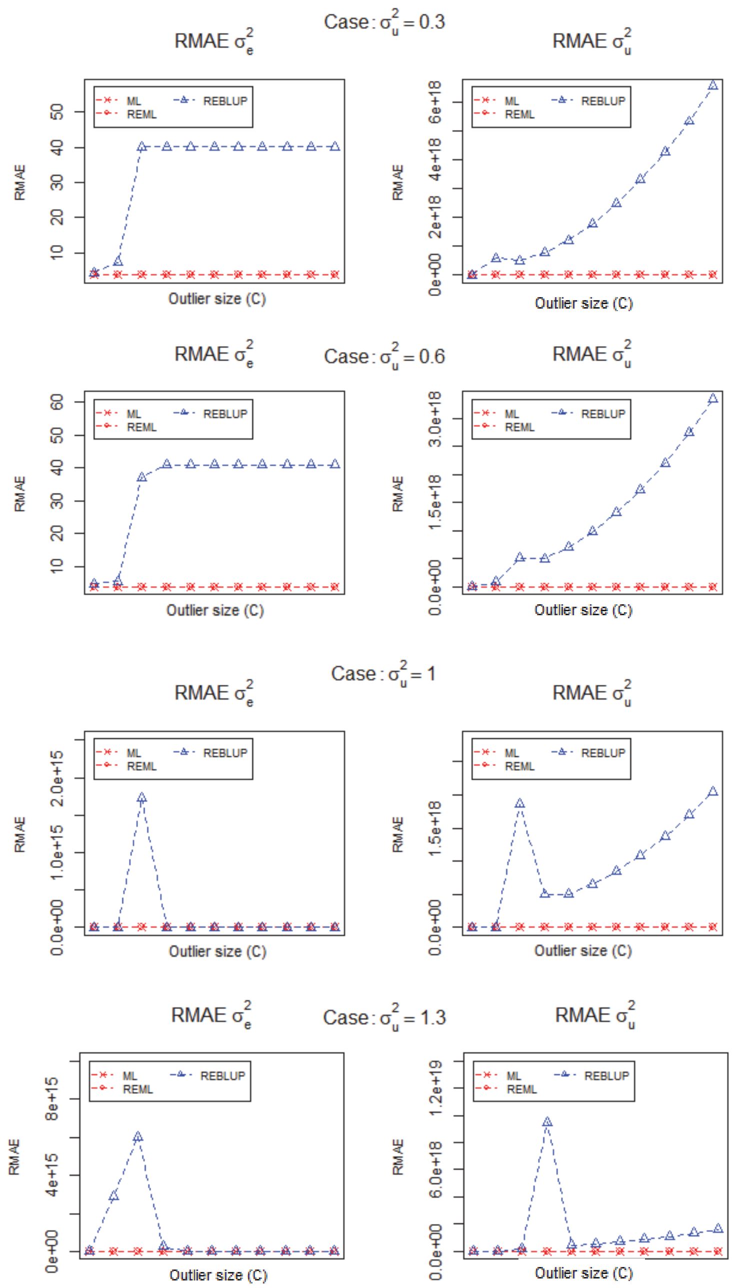


Figure A4. RMAE of ML, REML and REBLUP estimators under Scenario 3 (one outlying area)



Property 1:

If **A** is a non-singular and symmetric matrix of order *n* and has elements which are functions of a scalar w, then

$$\frac{\delta \ln(|\mathbf{A}|)}{\delta \mathrm{w}} = \mathrm{trace}\left\{\mathbf{A}^{-1}\right\}\frac{\delta \mathbf{A}}{\delta \mathrm{w}} .$$

Property 2:

If **A** is a non-singular matrix of order *n* and has elements which are functions of a scalar w, then

$$\frac{\delta \mathbf{A}^{-1}}{\delta \mathrm{w}} = -\mathbf{A}^{-1}\frac{\delta \mathbf{A}}{\delta \mathrm{w}}\mathbf{A}^{-1} .$$

Property 3:

Let **A** be a symmetric matrix of order *n* and **x** a vector of size *n*. If *f(x)* has a quadratic form, $f(x) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$, then $f'(x) = 2\mathbf{A}\mathbf{x}$.

References

BATTESE, G. E. – HARTER, R. M. – FULLER, W. A. [1988]: An error-components model for prediction of county crop areas using survey and satellite data. *Journal of the American Statistical Association.* Vol. 83. No. 401. pp. 28–36.

FELLNER, W. H. [1986]: Robust estimation of variance components. *Technometrics*. Vol. 28. No. 1. pp. 51–60. http://dx.doi.org/10.1080/00401706.1986.10488097

GERVINI, D. – YOHAI, V. J. [1998]: Robust estimation of variance components. *The Canadian Journal of Statistics*. Vol. 26. No. 3. pp. 419–430. http://dx.doi.org/10.2307/3315766

GURKA, M. – LLOYD, J. [2007]: Mixed models. In: *Rao, C. R. – Miller, J. P. – Rao, D. C*. (eds.): *Essential Statistical Methods for Medical Statistics*. Elsevier. Philadelphia. pp. 146–173. https://doi.org/10.1016/C2009-0-62044-4

HENDERSON, C. R. [1975]: Best linear unbiased estimation and prediction under a selection model. *Biometrics*. Vol. 31. No. 2. pp. 423–447.

HUBER, P. J. [1964]: Robust estimation of a location parameter. *The Annals of Mathematical Statistics.* Vol. 35. No. 1. pp. 73–101. http://dx.doi.org/10.1214/aoms/1177703732

JIANG, J. [1996]: REML estimation: asymptotic behavior and related topics. *The Annals of Statistics*. Vol. 24. No. 1. pp. 255–286. http://dx.doi.org/10.1214/aos/1033066209

MCCULLOCH, CH. – SEARLE, S. [2001]: *Generalized, Linear and Mixed Models*. Wiley Series in Probability and Statistics. John Wiley and Sons Ltd. Hoboken. http://dx.doi.org/10.1002/0471722073

MOLINA, I. – PEÑA, D. – PÉREZ, B. [2009]: *Robust Estimation in Linear Regression Models with Fixed Effects*. Universidad Carlos III De Madrid Working papers. No. 09-88 (27). https://core.ac.uk/download/pdf/30041693.pdf

PÉREZ, B. – MOLINA, I. – THIELER, A. – FRIED, R. – PEÑA, D. [2017]: Fast and robust estimators of variance components in the nested error model. *Statistics and Computing*. Vol. 27. No. 6. pp. 1655–1675. http://dx.doi.org/10.1007/s11222-016-9710-x

PÉREZ, B. – PEÑA, D. – MOLINA, I. [2011]: Robust Henderson III estimators of variance components in the nested error model. In: *Pardo, L. – Balakrishnan, N. – Gil, M. A.* (eds.): *Modern Mathematical Tools and Techniques in Capturing Complexity*. Springer. Berlin. pp. 329–339.

PÉREZ, B. [2011]: *Robust Estimation and Outlier Detection in Linear Models for Grouped Data*. Doctoral Thesis. Universidad Carlos III de Madrid. Madrid.

RICHARDSON, A. M. – WELSH, A. H. [1997]: Robust restricted maximum likelihood in mixed linear models. *Biometrics*. Vol. 51. No. 4. pp. 1429–1439. http://dx.doi.org/10.2307/2533273

ROCKE, D. M. [1983]: Robust statistical analysis of interlaboratory studies. *Biometrika*. Vol. 70. No. 2. pp. 421–431. http://dx.doi.org/10.1093/biomet/70.2.421

ROCKE, D. M. [1991]: Robustness and balance in the mixed model. *Biometrics*. Vol. 47. No. 1. pp. 303–309. http://dx.doi.org/10.2307/2532515

SINHA, S. K. – RAO, J. N. K. [2009]: Robust small area estimation. *The Canadian Journal of Statistics*. Vol. 37. No. 3. pp. 381–399. http://dx.doi.org/10.1002/cjs.10029

STAHEL, W. A. – WELSH, A. [1997]: Approaches to robust estimation in the simplest variance components model. *Journal of Statistical Planning and Inference*. Vol. 57. No. 2. pp. 295–319. http://dx.doi.org/10.1016/S0378-3758(96)00050-X