



Empirical study of computational techniques used for parameters' estimation in multivariate linear mixed effects models

Sèwanou H. Honfo^{1,*}, Romain Glèlè Kakaï¹

¹Laboratoire de Biomathématiques et d'Estimations Forêstières (LABEF), Faculty of Agronomic Sciences (FSA), University of Abomey-Calavi (UAC), Campus, Abomey-Calavi, 04 BP 1525 Cotonou

Received on August 09, 2019; Accepted on September 02, 2019

Copyright © 2019, The African Journal of Applied Statistics (AJAS) and The Statistics and Probability African Society (SPAS). All rights reserved

Abstract. An empirical study was conducted to assess the performance of some computational techniques used in multivariate linear mixed effects models (MLMM). Performances of four computational techniques were compared based on their accuracy, relative bias, relative efficiency and computing time according to sample size, between-variables and within-subjects correlations. The accuracy, relative bias, relative efficiency and computing time in estimating MLMM parameters varied significantly according to the computational techniques, sample size and correlations. Further theoretical developments are required to improve the accuracy of the four techniques in estimating variances' components.

Key words: multivariate linear mixed models; simulation; estimation accuracy; estimation relative bias; efficiency.

AMS 2010 Mathematics Subject Classification Objects : 81T80; 62H12; 68Q25; 65Y20

*Corresponding author Sèwanou H. Honfo: honfosewanou@gmail.com
Romain Glèlè Kakaï : glele.romain@gmail.com

Résumé. (Abstract in French) Une étude empirique a été conduite pour évaluer la performance de certaines méthodes d'estimation utilisées pour l'ajustement des modèles multivariés linéaires à effets mixtes (MLMM). Quatre méthodes d'estimation majeures ont été comparées sur la base de leur précision, biais relatif, efficacité relative et temps de calcul en fonction de la taille d'échantillon, et les corrélations entre variables dépendantes et entre les sujets. Les critères de performance d'estimation des paramètres des MLMMs variaient significativement suivant les méthodes d'estimation, la taille d'échantillon et les corrélations. Des études plus approfondies sont requises pour améliorer la précision des méthodes considérées à estimer les variances dues aux effets aléatoires et aux résidus du modèle.

1. Introduction

Applied research sciences such as socio-economy, ecology, medicine, biology and botany often aim to model experimental data collected over time repeatedly (Verbeke *et al.*, 2014). Multilevel, clustered and longitudinal data are widely used in these fields because they allow assessing changes over time as well as identifying factors that affect changes' patterns (An *et al.*, 2013). Linear Mixed effects Model (LMM) (Laird and Ware , 1982) and its extensions (Jenrich and Schluchter , 1986; Laird *et al.*, 1987; Lindstrom and Bates , 1988; Chi and Reinsel , 1989) are popular models for experimental data.

Research studies in applied sciences often consider many variables of interest to explain phenomena. For this issue, Multivariate LMM (MLMM) methods have been developed as a generalization frame of LMM to handle this frequent situation. MLMMs have the particularity to model unbalanced multivariate longitudinal data and therefore have become the most appropriate modeling methods for continuous longitudinal data with multiple response variables (Wang and Fan , 2010).

Since its first development (Reinsel , 1982, 1984), various MLMM computational techniques are developed (Shah *et al.*, 1997; Reinsel , 1984; Sammel *et al.*, 1999; Roy and Lin , 2000; Reeves and MacKenzie , 1998; Schafer and Yucel , 2002; Fieuws and Verbeke , 2006; Wang and Fan , 2010; Adjakossa , 2017). Several studies compared some of them and indicated significant difference in terms of parameters' estimation (Mingers , 1989; Tekwe *et al.*, 2004) and existing computational issues. One of properties of reliable computation method is its efficiency both for estimation accuracy and convergence speed (Schafer and Yucel , 2002; Fieuws *et al.*, 2007; Wang and Fan , 2010). For instance, the approach of Reinsel (1982, 1984) is only applicable for balanced and complete design in which all subjects are measured at the same time points (no missing observations). Computational techniques based on the Expectation-Maximization (EM) algorithm (Shah *et al.*, 1997) and hybrid EM-Fisher scoring algorithm (Schafer and Yucel , 2002) present weaknesses in estimating parameters for large variance data and important number of clusters, respectively.

In the last decade, new computational approaches were developed but a shadow persists on their efficiency compared to old computational techniques in the face of major factors influencing parameters' estimation in the context of MLMs: sample size and autocorrelation between-variables and within-subjects. Ideally, new approaches should take into account the weaknesses of old ones to improve the accuracy in parameters estimation so that users of MLMs can make reliable decisions and in general make advance statistical methods. This is not often the case and it is important to evaluate the real practical progress that the new approaches offer in terms of accuracy in estimating MLMs parameters. Therefore, this study aims to assess the accuracy of parameters' estimation according to variations of sample sizes and correlations between-variables and within subjects through an empirical study.

We assume that (i) the performance of MLM computational techniques used for fitting multivariate (longitudinal) data varies according to sample sizes and correlations (Fieuw and Verbeke, 2006; Wang and Fan, 2010; Adjakossa et al., 2016) and (ii) new computational techniques improve the precision in estimating parameters in the frame of MLMs.

2. Model formulation

Let consider a longitudinal study with n_j subjects ($j = 1, \dots, p$). The study consists in observing p parameters over time on each subject i ($i = 1, \dots, n_j$). Let $\mathbf{Y}_i = (\mathbf{y}_{i1}, \dots, \mathbf{y}_{ip})^T$ be an $k_i \times p$ matrix of response variables (associated to the observed parameters) for a given subject i , where each $\mathbf{y}_{ij} = (\mathbf{y}_{ij1}; \dots; \mathbf{y}_{ijk_i})^T$ is a $k_i \times 1$ vector of the j^{th} parameters measured at scheduled time points $t = 1, \dots, k_i$. Without loss of generality, let us consider a bivariate linear mixed effects model (Adjakossa, 2017) defined as

$$\begin{bmatrix} \mathbf{Y}_{1i} \\ \mathbf{Y}_{2i} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_{1i}\beta_1 + \mathbf{Z}_{1i}\gamma_{1i} + \epsilon_{1i} \\ \mathbf{X}_{2i}\beta_2 + \mathbf{Z}_{2i}\gamma_{2i} + \epsilon_{2i} \end{bmatrix}, \quad (1)$$

where \mathbf{X}_{1i} and \mathbf{X}_{2i} are $k_i \times q_1$ and $k_i \times q_2$ matrices of fixed covariates related to \mathbf{Y}_{1i} and \mathbf{Y}_{2i} , respectively; \mathbf{Z}_{1i} and \mathbf{Z}_{2i} are $k_i \times r_1$ and $k_i \times r_2$ matrices of random covariates related to \mathbf{Y}_{1i} and \mathbf{Y}_{2i} , respectively; in a bivariate case; β_1 and β_2 are fixed effects vectors related to \mathbf{X}_1 and \mathbf{X}_2 , respectively.

$$\gamma_i = \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} \sim N \left(\mathbf{0}; \bar{\Gamma} = \begin{pmatrix} \bar{\Gamma}_1 & \bar{\Gamma}_{12} \\ \bar{\Gamma}_{21} & \bar{\Gamma}_2 \end{pmatrix} \right) \quad (2)$$

where γ_1 and γ_2 are (r_1, r_2) -dimensional vectors of random effects associated with the random covariates \mathbf{Z}_1 and \mathbf{Z}_2 respectively, with $\bar{\Gamma}$, the variance-covariance matrix of random effects considered as follow,

$$\bar{\Gamma}_1 = \begin{pmatrix} \eta_1^2 & \rho_\eta \eta_1 \eta_2 \\ \rho_\eta \eta_2 \eta_1 & \eta_2^2 \end{pmatrix}, \bar{\Gamma}_2 = \begin{pmatrix} \tau_1^2 & \rho_\tau \tau_1 \tau_2 \\ \rho_\tau \tau_2 \tau_1 & \tau_2^2 \end{pmatrix}, \bar{\Gamma}_{12} = \begin{pmatrix} \rho_{\eta_1 \tau_1} \eta_1 \tau_1 & \rho_{\eta_1 \tau_2} \eta_1 \tau_2 \\ \rho_{\eta_2 \tau_1} \eta_2 \tau_1 & \rho_{\eta_2 \tau_2} \eta_2 \tau_2 \end{pmatrix} \text{ so that}$$

$$V(\gamma_1) = \bar{\Gamma}_1, V(\gamma_2) = \bar{\Gamma}_2, \text{ and } Cov(\gamma_1, \gamma_2) = \bar{\Gamma}_{12}.$$

$$\varepsilon_i = \begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix} \sim N \left(\mathbf{0}; \bar{\Sigma} = \begin{pmatrix} \sigma_1^2 \mathbf{I}_{k_{1i}} & \delta \sigma_1 \sigma_2 \mathbf{I}_{k_{1i}} \\ \delta \sigma_1 \sigma_2 \mathbf{I}_{k_{2i}} & \sigma_2^2 \mathbf{I}_{k_{2i}} \end{pmatrix} \right) \quad (3)$$

where ε_i represents the vector of models' residuals, $\bar{\Sigma}$, the variance-covariance matrix of residuals, σ_1^2 and σ_2^2 are residuals variances related to \mathbf{Y}_1 and \mathbf{Y}_2 , δ is a constant correlation between \mathbf{Y}_{1i} and \mathbf{Y}_{2i} at each time point.

3. MLMM computational techniques considered

3.1. MLMM via EM Algorithm

The EM algorithm is a computational process that consists in estimating parameters of a model through two steps: E-step (creation of expectation function of the log-likelihood) and M-step (computation of parameters that maximize the expected log-likelihood) at each iteration until convergence (Dempster *et al.*, 1977). Shah *et al.* (1997) proposed ML and REML estimations of parameters in MLMMs with possible missing data through EM algorithm. Here, we only presented the ML estimation method for complete observations (Shah *et al.*, 1997).

E-step: Let θ be the vector of unknown parameters in Σ ($\Sigma = \bar{\Sigma}$ in (3)) and D ($D = \bar{D}$ in (2)) and let $\theta^{(\tau)}$ denotes their estimates at the end of the τ th iteration.

$$E[\beta^{(\tau)} | \theta^{(\tau)}] = \left[\sum_{i=1}^n \mathbf{X}_i^T \mathbf{P}_i^{(\tau)} \mathbf{X}_i \right]^{-1} \sum_{i=1}^n \mathbf{X}_i^T \mathbf{P}_i^{(\tau)} \mathbf{y}_i \quad (4)$$

where $\mathbf{y}_i = \begin{pmatrix} \mathbf{Y}_{1i} \\ \mathbf{Y}_{2i} \end{pmatrix}$, $\mathbf{X}_i = \begin{pmatrix} \mathbf{X}_{1i} \\ \mathbf{X}_{2i} \end{pmatrix}$ and $\mathbf{Z}_i = \begin{pmatrix} \mathbf{Z}_{1i} \\ \mathbf{Z}_{2i} \end{pmatrix}$ are respectively $p_{k_i} \times q$, $q = (q_1, q_2)^T$

and $p_{k_i} \times r$, $r = (r_1, r_2)^T$ known fixed and random covariates matrices, $\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$,

$\mathbf{P}_i^{(\tau)} = \mathbf{V}_i^{(\tau)-1}$ and $\mathbf{V}_i = var(\mathbf{y}_i) = [\mathbf{Z}_i \mathbf{D} \mathbf{Z}_i^T + \Sigma \otimes \mathbf{I}_i]$.

The joint density of \mathbf{y}_i , γ_i , and ε_i is used to obtain the conditional expectations of the sufficient statistics. Considering $\nu_i = \mathbf{y}_i - \mathbf{X}_i \beta$, the expectations of the i th term of the sufficient statistics for given β and θ are defined as (Shah *et al.*, 1997):

$$E[\gamma_i \gamma_i^T | \mathbf{y}_i] = E[\gamma_i | \mathbf{y}_i] E[\gamma_i | \mathbf{y}_i]^T + V[\gamma_i | \mathbf{y}_i] \quad (5)$$

and

$$E[\varepsilon_{ij}^T \varepsilon_{ik} | \mathbf{y}_i] = E[\varepsilon_{ij} | \mathbf{y}_i]^T E[\varepsilon_{ik} | \mathbf{y}_i] + tr[cov(\varepsilon_{ij}, \varepsilon_{ik}) | \mathbf{y}_i], \quad j, k = 1, 2. \quad (6)$$

where $E[\gamma_i | \mathbf{y}_i] = \mathbf{DZ}_i^T \mathbf{P}_i \boldsymbol{\nu}_i$; $V[\gamma_i | \mathbf{y}_i] = \mathbf{D} - \mathbf{DZ}_i^T \mathbf{P}_i \mathbf{Z}_i \mathbf{D}$; $E[\mathbf{e}_i | \mathbf{y}_i] = \boldsymbol{\nu}_i - \mathbf{Z}_i E[\gamma_i | \mathbf{y}_i]$; and $V[\mathbf{e}_i | \mathbf{y}_i] = \boldsymbol{\Sigma} \otimes \mathbf{I}_i - (\boldsymbol{\Sigma} \otimes \mathbf{I}_i) \mathbf{P}_i (\boldsymbol{\Sigma} \otimes \mathbf{I}_i)$, with $\mathbf{e}_i = \text{vec}(\boldsymbol{\varepsilon}) = [\boldsymbol{\varepsilon}_{1i}^T, \boldsymbol{\varepsilon}_{2i}^T]^T$.

M-step: This step allows obtaining $\boldsymbol{\Sigma}^{\tau+1}$ and $\mathbf{D}^{\tau+1}$ by using:

$$\mathbf{D}^{\tau+1} = \frac{\sum_{i=1}^n [E[(\gamma_i)(\gamma_i)^T | \mathbf{y}_i, \boldsymbol{\theta}^\tau, \boldsymbol{\beta}^\tau]]}{N} \quad (7)$$

and

$$\boldsymbol{\Sigma}^{\tau+1} = \left[\sum_{i=1}^n k_i \right]^{-1} \left[\sum_{i=1}^n [E[(\boldsymbol{\varepsilon}_{ij})^T (\boldsymbol{\varepsilon}_{ik}) | \mathbf{y}_i, \boldsymbol{\theta}^\tau, \boldsymbol{\beta}^\tau]] \right], \quad j, k = 1, 2. \quad (8)$$

3.2. MLMM via hybrid of EM and Fisher scoring Algorithm

Schafer and Yucel (2002) have developed a set of methods to model multivariate longitudinal or clustered data with the particularity of handling missing observations. One of these computational techniques is the EM-Fisher scoring algorithm (FS-EM) that uses Fisher scoring technique in the M-step of EM. FS-EM already exists in R software as package, namely *mlmm* (Schafer and Yucel, 2002).

E-step: It computes the expectation of the complete-data log-likelihood function which consists in computing expectations of two sufficient statistics \mathbf{y}_i and $\mathbf{y}_i \mathbf{y}_i^T$.

$$E(\mathbf{y}_i | \mathbf{y}_{i(obs)}) = E(E(\mathbf{y}_i | \mathbf{y}_{i(obs)}, \gamma_i) | \mathbf{y}_{i(obs)}) \quad (9)$$

$$E(\mathbf{y}_{ijk} \mathbf{y}_{ij'k'} | \mathbf{y}_{i(obs)}) = \begin{cases} \mathbf{y}_{ijk} \mathbf{y}_{ij'k'} & \text{if } \mathbf{y}_{ijk}^{(observed)} \text{ and } \mathbf{y}_{ij'k'}^{(observed)} \\ \mathbf{y}_{ijk} E(\mathbf{y}_{ij'k'} | \mathbf{y}_{i(obs)}) & \text{if } \mathbf{y}_{ijk}^{(observed)} \text{ and } \mathbf{y}_{ij'k'}^{(missing)} \\ E(\mathbf{y}_{ijk} | \mathbf{y}_{i(obs)}) E(\mathbf{y}_{ij'k'} | \mathbf{y}_{i(obs)}) + & \\ \text{cov}(\mathbf{y}_{ijk}, \mathbf{y}_{ij'k'} | \mathbf{y}_{i(obs)}) & \text{if } \mathbf{y}_{ijk}^{(missing)} \text{ and } \mathbf{y}_{ij'k'}^{(missing)} \end{cases} \quad (10)$$

where,

$$\text{cov}(\mathbf{y}_{ijk}, \mathbf{y}_{ij'k'} | \mathbf{y}_{i(obs)}) = \begin{cases} \text{cov}(A_{ijk}, A_{ij'k'} | \mathbf{y}_{i(obs)}) + [\boldsymbol{\Sigma}_{22.1}]_{kk'} & \text{if } j = j' \\ \text{cov}(A_{ijk}, A_{ij'k'} | \mathbf{y}_{i(obs)}) & \text{if } j \neq j' \end{cases},$$

where $A_{ijk} = E(\mathbf{y}_{ijk} | \gamma_i, \mathbf{y}_{i(obs)})$, \mathbf{y}_{ijk} denotes the k th element of the j th row of \mathbf{y}_i .

M-step: It consists in computing first derivatives of the log-likelihood expectation given $\mathbf{Y}_{(mis)}$ with respect to elements of parameters θ (11). Then, parameters are updated using the Fisher scoring procedure.

$$\begin{cases} \frac{\partial \ell_e}{\partial \text{vec}(\beta)} = - (\sum_{i=1}^n (\mathbf{I}_r \otimes \mathbf{X}_i)^T W_i (\mathbf{I}_r \otimes \mathbf{X}_i)) \text{vec}(\beta - \tilde{\beta}), \\ \frac{\partial \ell_e}{\partial \omega_j} = \frac{1}{2} \sum_{i=1}^n \text{tr}(\bar{\Gamma} - \mathbf{U}_i - (\bar{\Sigma}^{-1} \otimes \mathbf{Z}_i^T \mathbf{Z}_i) \mathbf{U}_i \mathbf{T}_i \mathbf{U}_i (\bar{\Sigma}^{-1} \otimes \mathbf{Z}_i^T \mathbf{Z}_i)) \mathbf{G}_j, \\ \frac{\partial \ell_e}{\partial \sigma_l} = \frac{1}{2} \sum_{i=1}^n \text{tr}(n_i \bar{\Sigma} \mathbf{F}_l - (\mathbf{F}_l \otimes \mathbf{Z}_i^T \mathbf{Z}_i) \mathbf{U}_i - W_i (\bar{\Sigma} \mathbf{F}_l \bar{\Sigma} \otimes \mathbf{I}_{n_i}) W_i \mathbf{T}_i) \end{cases} \quad (11)$$

where

$$\begin{aligned} W_i^{-1} &= (\mathbf{I}_r \otimes \mathbf{Z}_i) \bar{\Gamma} (\mathbf{I}_r \otimes \mathbf{Z}_i)^T + (\bar{\Sigma} \otimes \mathbf{I}_{n_i}), \\ \mathbf{U}_i &= (\bar{\Sigma}^{-1} + (\bar{\Sigma} \otimes \mathbf{Z}_i^T \mathbf{Z}_i))^{-1}, \\ \text{vec}(\tilde{\beta}) &= \Lambda \sum_{i=1}^n (\mathbf{I}_r \otimes \mathbf{X}_i)^T W_i E(\text{vec}(\mathbf{y}_i) \mid \theta, \mathbf{y}_{i(obs)}), \\ \Lambda^{-1} &= \sum_{i=1}^n (\mathbf{I}_r \otimes \mathbf{X}_i)^T W_i (\mathbf{I}_r \otimes \mathbf{X}_i), \\ \mathbf{T}_i &= E \{ \text{vec}(\mathbf{y}_i - \mathbf{X}_i \beta) \text{vec}(\mathbf{y}_i - \mathbf{X}_i \beta)^T \mid \mathbf{y}_{i(obs)}, \theta \}. \end{aligned}$$

3.3. MLMM via ECM Algorithm

The ECM algorithm (Meng and Rubin, 1993) is a general extension of EM in which the M-step is replaced by many computationally simpler conditional maximization (CM) steps. Wang and Fan (2010) have proposed an efficient implementation of ECM for fitting MLMMs.

E-step: It consists in computing $\Upsilon(\theta \mid \hat{\theta}^{(\tau)})$, the conditional expectation of the complete data log-likelihood function given the observed data \mathbf{y}_i and the current estimates $\hat{\theta}^{(\tau)}$.

$$\Upsilon(\theta \mid \hat{\theta}^{(\tau)}) = -\frac{1}{2} \sum_{i=1}^n \left\{ \log \mid \nabla \otimes \mathbf{C}_i \mid + \log \mid \bar{\Gamma} \mid + \text{tr} \left((\nabla \otimes \mathbf{C}_i)^{-1} \hat{\mathbf{E}}_i^{(\tau)} + \bar{\Gamma}^{-1} \wp^{(\tau)} \right) \right\} \quad (12)$$

where $\nabla \otimes \mathbf{C}_i = \bar{\Sigma}$ in (3), with $\nabla = [\sigma_{jj'}^2]$ being a $p \times p$ unstructured matrix describing the variance-covariance among response variables, and $\mathbf{C}_i = \mathbf{C}_i(\phi) = [\varrho_{|t-t'|}(\phi)]$, for $t, t' = 1, \dots, k_i$, is a $k_i \times k_i$ structured AR(p)-process matrix which allows addressing the autocorrelation among k_i occasions on each response variable. Note that ϱ can

be obtained through the Yule-Walker equation (Box *et al.*, 1994; Wang and Fan , 2010; Box *et al.*, 2015),

$$\varrho_s(\phi) = \varrho_s = \phi_1 \varrho_{s-1} + \dots + \phi_p \varrho_{s-p}, \quad \varrho_0 = 1, \quad (s = 0, \dots, k_i - 1), \quad \phi = (\phi_1, \dots, \phi_p)^T.$$

Then

$$\hat{\mathbf{E}}_i^{(\tau)} = \hat{\mathbf{E}}_i^{(\tau)}(\beta) = E(\varepsilon_i \varepsilon_i^T | \mathbf{y}_i, \hat{\boldsymbol{\theta}}^{(\tau)}) = \hat{\varepsilon}_i^{(\tau)} \hat{\varepsilon}_i^{(\tau)T} + \hat{\mathbf{V}}_{\varepsilon_i}^{(\tau)},$$

where

$$\hat{\varepsilon}_i^{(\tau)} = E(\varepsilon_i | \mathbf{y}_i, \hat{\boldsymbol{\theta}}^{(\tau)}) = \begin{bmatrix} \mathbf{Y}_1 - \mathbf{X}_1 \beta_1 \\ \mathbf{Y}_2 - \mathbf{X}_2 \beta_2 \end{bmatrix} - (\mathbf{I}_p \otimes \mathbf{Z}_i) \gamma_i^{(\tau)},$$

$$\hat{\mathbf{V}}_{\varepsilon_i}^{(\tau)} = cov(\varepsilon_i | \mathbf{y}_i, \hat{\boldsymbol{\theta}}^{(\tau)}) = \left[\mathbf{I}_{k_i p} - \left(\hat{\mathbf{V}}^{(\tau)} \otimes \hat{\mathbf{C}}_i^{(\tau)} \right) \hat{\boldsymbol{\Lambda}}_i^{(\tau)-1} \right] \left(\hat{\mathbf{V}}^{(\tau)} \otimes \hat{\mathbf{C}}_i^{(\tau)} \right),$$

$$\hat{\boldsymbol{\Lambda}}_i^{(\tau)} = (\mathbf{I}_p \otimes \mathbf{Z}_i) \hat{\boldsymbol{\Gamma}}^{(\tau)} (\mathbf{I}_p \otimes \mathbf{Z}_i)^T + \left(\hat{\mathbf{V}}^{(\tau)} \otimes \hat{\mathbf{C}}_i^{(\tau)} \right), \quad \hat{\mathbf{C}}_i^{(\tau)} = \mathbf{C}_i(\hat{\phi}^{(\tau)});$$

and

$$\wp^{(\tau)} = E(\gamma_i \gamma_i^T | \mathbf{y}_i, \hat{\boldsymbol{\theta}}^{(\tau)}) = \hat{\gamma}_i^{(\tau)} \hat{\gamma}_i^{(\tau)T} + \hat{\mathbf{V}}_{\gamma_i}^{(\tau)},$$

where

$$\hat{\gamma}_i^{(\tau)} = E(\gamma_i | \mathbf{y}_i, \hat{\boldsymbol{\theta}}^{(\tau)}) = \hat{\boldsymbol{\Gamma}}^{(\tau)} (\mathbf{I}_p \otimes \mathbf{Z}_i)^T \hat{\boldsymbol{\Lambda}}_i^{(\tau)-1} \begin{bmatrix} \mathbf{Y}_1 - \mathbf{X}_1 \hat{\beta}_1^{(\tau)} \\ \mathbf{Y}_2 - \mathbf{X}_2 \hat{\beta}_2^{(\tau)} \end{bmatrix},$$

$$\hat{\mathbf{V}}_{\gamma_i}^{(\tau)} = cov(\gamma_i | \mathbf{y}_i, \hat{\boldsymbol{\theta}}^{(\tau)}) = \left[\hat{\boldsymbol{\Gamma}}^{(\tau)-1} + (\mathbf{I}_p \otimes \mathbf{Z}_i)^T \left(\hat{\mathbf{V}}^{(\tau)} \otimes \hat{\mathbf{C}}_i^{(\tau)} \right)^{-1} (\mathbf{I}_p \otimes \mathbf{Z}_i) \right]^{-1}.$$

M-step: The M-step is constituted of 4 CM-steps, each of which maximizes (12) over θ but conditioned on some vector functions of θ being estimated at its previous step, proceed as follows:

CM-step 1 : Update $\hat{\beta}^{(\tau)}$ by maximizing (12) over β , which leads to

$$\hat{\beta}^{(\tau+1)} = \left[\sum_{i=1}^n (\mathbf{I}_p \otimes \mathbf{X}_i)^T \left(\hat{\mathbf{V}}^{(\tau)} \otimes \hat{\mathbf{C}}_i^{(\tau)} \right)^{-1} (\mathbf{I}_p \otimes \mathbf{X}_i) \right]^{-1} \left[\sum_{i=1}^n (\mathbf{I}_p \otimes \mathbf{X}_i)^T \left(\hat{\mathbf{V}}^{(\tau)} \otimes \hat{\mathbf{C}}_i^{(\tau)} \right)^{-1} \left(\mathbf{y}_i - (\mathbf{I}_p \otimes \mathbf{Z}_i) \gamma_i^{(\tau)} \right) \right]$$

CM-step 2 : Update $\hat{\Gamma}^{(\tau+1)}$ by maximizing (12) over $\bar{\Gamma}$, which gives

$$\hat{\Gamma}^{(\tau+1)} = \frac{1}{n} \sum_{i=1}^n \wp^{(\tau)}.$$

CM-step 3 : Define $\nabla^{-1} = [\sigma^{2jj'}]$ and $\nabla = [\sigma_{jj'}^2]$, for $j, j' = 1, \dots, p$. Differentiating (12)

with respect to $\sigma^{2jj'}$ and setting it to zero provide,

$$\sigma_{jj'}^{2(\hat{\tau}+1)} = \begin{cases} (\sum_{i=1}^n k_i)^{-1} \sum_{i=1}^n \text{tr} \left(\hat{C}_i^{(\tau)-1} \hat{E}_{ijj'}^{(\tau)} (\hat{\beta}^{(\tau+1)}) \right) & \text{for } j = j' \\ (2 \sum_{i=1}^n k_i)^{-1} \sum_{i=1}^n \text{tr} \left(\hat{C}_i^{(\tau)-1} (\hat{E}_{ijj'}^{(\tau)} (\hat{\beta}^{(\tau+1)}) + \hat{E}_{ijj'}^{(\tau)} (\hat{\beta}^{(\tau+1)})) \right) & \text{for } j \neq j' \end{cases}$$

$\hat{E}_{ijj'}^{(\tau)} = \hat{E}_{ijj'}^{(\tau)}(\beta) = E(\varepsilon_{ij} \varepsilon_{ij'}^T \mid \mathbf{y}_i, \hat{\theta}^{(\tau)}) = \hat{\varepsilon}_{ij}^{(\tau)} \hat{\varepsilon}_{ij'}^{(\tau)T} + \mathbf{Z}_i \hat{\mathbf{V}}_{\gamma_{ijj'}}^{(\tau)} \mathbf{Z}_i^T$ is a square submatrix of order k_i , where $\hat{\varepsilon}_{ij}^{(\tau)} = \mathbf{y}_{ij} - \mathbf{X}_i \beta_j - \mathbf{Z}_i \hat{\gamma}_{ij}^{(\tau)}$, $\hat{\varepsilon}_{ij'}^{(\tau)} = \mathbf{y}_{ij'} - \mathbf{X}_i \beta_{j'} - \mathbf{Z}_i \hat{\gamma}_{ij'}^{(\tau)}$ with $\hat{\gamma}_{ij'}^{(\tau)}$ being a $r \times 1$ subvector consisting of the $((j-1)r+1)$ th to the (jr) th entries of $\hat{\gamma}_i^{(\tau)}$, and $\hat{\mathbf{V}}_{\gamma_{ijj'}}^{(\tau)}$ being a $r \times r$ submatrix consisting of $((j-1)r+1)$ th to the (jr) th rows and the $((l-1)r+1)$ th to the (lr) th columns of $\hat{\mathbf{V}}_{\gamma_i}^{(\tau)}$, for $j, j' = 1, \dots, p$, $r = r_1 + r_2$.

CM-step 4 : Update $\hat{\pi}^{(\tau)}$ by maximizing (12) over π , which yields

$$\hat{\pi}^{(\tau+1)} = \underset{\pi \in [-1, 1]^p}{\text{argmax}} \left\{ \sum_{i=1}^n \left[p \log | \mathbf{C}_i^{-1}(\pi) | - \text{tr} \left((\nabla^{-1} \otimes \mathbf{C}_i^{-1}(\pi)) \hat{E}_i^{(\tau)} \right) \right] \right\}.$$

Once $\hat{\pi}^{(\tau+1)}$ is obtained, $\hat{\phi}^{(\tau+1)}$ can be computed through a transformation according to Barndorff-Nielsen and Schou (1973) (details can be found in Wang and Fan (2010)).

3.4. MLMM via CMLME Algorithm

Consistent Estimates for the Multivariate Linear Mixed-Effects Models (CMLME), developed by Adjakossa (2017), is a recent ML (and REML) method to estimate parameters in the frame of MLMMs. The computational technique is mainly characterized by a profiling of the model's deviance and a Cholesky's factorization of the random effect covariance matrix. The method's formulation (Adjakossa, 2017) is limited to between-random effects correlations and does not include marginal residuals correlations. In the frame of this study we focused on ML estimation procedure.

Parameters to be estimated

Considering $\bar{\Gamma}_1$ and $\bar{\Gamma}_2$ in (2) as positive semidefinite matrix, they can be easily expressed in terms of relative covariance factors, Λ_{θ_1} and Λ_{θ_2} which are $r_1 \times r_1$ and $r_2 \times r_2$ matrices, respectively. Λ_{θ_1} and Λ_{θ_2} are block diagonal matrices of which

each diagonal element is a lower triangular matrix whose nonzero entries are the components of vectors θ_1 and θ_2 , respectively. θ_1 and θ_2 provide the symmetric $r_1 \times r_1$ and $r_2 \times r_2$ variance-covariance matrices $\bar{\Gamma}_1$ and $\bar{\Gamma}_2$, according to (Adjakossa , 2017)

$$\bar{\Gamma}_1 = \sigma_1^2 \Lambda_{\theta_1} \Lambda_{\theta_1}^T. \tag{13}$$

$$\bar{\Gamma}_2 = \sigma_2^2 \Lambda_{\theta_2} \Lambda_{\theta_2}^T. \tag{14}$$

Then, the marginal random effects can be further expressed as

$$\begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} = \begin{pmatrix} \Lambda_{\theta_1} u_1 \\ \Lambda_{\theta_2} u_2 \end{pmatrix} \tag{15}$$

such that

$$u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \sim N(\mathbf{0}, \Sigma_u), \quad \Sigma_u = \begin{pmatrix} \sigma_1^2 \mathbf{I}_{r_1} & \sigma_1 \sigma_2 \rho \\ \sigma_1 \sigma_2 \rho^T & \sigma_2^2 \mathbf{I}_{r_2} \end{pmatrix} \tag{16}$$

In (13), ρ is a block diagonal matrix and u is a realization of a random vector z . The diagonal elements ρ of ρ are matrices containing correlations between γ_1 and γ_2 . Considering that each random effect is composed by two elements namely, intercept (I) and slope (S), ρ can be expressed as

$$\rho = \begin{pmatrix} \text{corr}(\gamma_1^I, \gamma_2^I) & \text{corr}(\gamma_1^I, \gamma_2^S) \\ \text{corr}(\gamma_1^S, \gamma_2^I) & \text{corr}(\gamma_1^S, \gamma_2^S) \end{pmatrix}; \quad \rho = \text{diag}(\rho, \dots, \rho). \tag{17}$$

Then, the bivariate LMM in (1) becomes

$$\begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 \beta_1 + \mathbf{Z}_1 \Lambda_{\theta_1} u_1 + \varepsilon_1 \\ \mathbf{X}_2 \beta_2 + \mathbf{Z}_2 \Lambda_{\theta_2} u_2 + \varepsilon_2 \end{bmatrix} \tag{18}$$

with u in (16) and ε in (3) where $\delta = 0$. Therefore, parameters to be estimated are $\beta_1, \beta_2, \sigma_1^2, \sigma_2^2, \theta_1, \theta_2, \rho$.

ML estimation via CMLME algorithm

Theorem 1 synthesizes the computational technique of CMLME (Adjakossa , 2017).

Theorem 1. Let $y = \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix}$ satisfying (18) where $\beta_1, \beta_2, \sigma_1^2, \sigma_2^2, \theta_1, \theta_2$ and ρ are parameters to be estimated; and $\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$, $\sigma = \begin{pmatrix} \sigma_1^2 \\ \sigma_2^2 \end{pmatrix}$, $\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}$. Denoting by $\mathbf{Y}_\sigma = \begin{pmatrix} \sqrt{\sigma_2^2} \mathbf{y}_1 \\ \sqrt{\sigma_1^2} \mathbf{y}_2 \end{pmatrix}$, $\mathbf{X}_\sigma = \begin{pmatrix} \sqrt{\sigma_2^2} \mathbf{X}_1 & \mathbf{0} \\ \mathbf{0} & \sqrt{\sigma_1^2} \mathbf{X}_2 \end{pmatrix}$, $\mathbf{Z}_{\sigma\theta} = \begin{pmatrix} \sqrt{\sigma_2^2} \mathbf{Z}_1 \Lambda_{\theta_1} & \mathbf{0} \\ \mathbf{0} & \sqrt{\sigma_1^2} \mathbf{Z}_2 \Lambda_{\theta_2} \end{pmatrix}$, and \hat{u} the conditional mean of z given y , the log-likelihood of β, σ, θ , and ρ given y is expressed as

$$-2\ell(\beta, \sigma, \theta, \rho | \mathbf{y}) = \frac{r(\hat{\beta}, \hat{u}) + \|\mathbf{R}_X(\beta - \hat{\beta})\|^2}{2} + (N - r_1 - r_2)\log(\sigma_1^2\sigma_2^2) + \frac{1}{2}\log(|\Sigma_u|) + \frac{1}{2}\log(\mathbf{L}) \quad (19)$$

where

$$\hat{u} = \left[\begin{array}{c} \mathbf{Z}_{\sigma\theta}^T \mathbf{Z}_{\sigma\theta} + \sqrt{\sigma_1^2\sigma_2^2} \Sigma_u^{-1} - \mathbf{Z}_{\sigma\theta}^T \mathbf{X}_\sigma [\mathbf{X}_\sigma^T \mathbf{X}_\sigma]^{-1} \mathbf{X}_\sigma^T \mathbf{Z}_{\sigma\theta} \\ \mathbf{Z}_{\sigma\theta}^T \mathbf{Y}_\sigma - \mathbf{Z}_{\sigma\theta}^T \mathbf{X}_\sigma [\mathbf{X}_\sigma^T \mathbf{X}_\sigma]^{-1} \mathbf{X}_\sigma^T \mathbf{Y}_\sigma \end{array} \right]^{-1} \quad (20)$$

$$\hat{\beta} = [\mathbf{X}_\sigma^T \mathbf{X}_\sigma]^{-1} [\mathbf{X}_\sigma^T \mathbf{Y}_\sigma - \mathbf{X}_\sigma^T \mathbf{Z}_{\sigma\theta} \hat{u}] \quad (21)$$

$$r(\hat{\beta}, \hat{u}) = \|\mathbf{Y}_\sigma - \mathbf{X}_\sigma \hat{\beta} - \mathbf{Z}_{\sigma\theta} \hat{u}\|^2 + \hat{u}^T \Sigma_u^{-1} \hat{u} \quad (22)$$

$$\mathbf{L} = \text{chol} \left(\mathbf{Z}_{\sigma\theta}^T \mathbf{Z}_{\sigma\theta} + \sqrt{\sigma_1^2\sigma_2^2} \Sigma_u^{-1} \right) \quad (23)$$

$$\mathbf{R}_X = \text{chol}(\mathbf{X}_\sigma^T \mathbf{X}_\sigma - \mathbf{R}_{ZX}^T \mathbf{R}_{ZX}), \quad \mathbf{R}_{ZX} = \mathbf{L}^{-1} \mathbf{Z}_{\sigma\theta}^T \mathbf{X}_\sigma. \quad (24)$$

with $\text{chol}(A)$, the Cholesky factorization of the symmetric positive definite square matrix A.

4. Simulation design

The bivariate LMM (1) was considered for data simulation and we assumed $n_1 = n_2$. Two computational factors were studied: sample size and correlations. The sample size was characterized by the number of subjects (n) and the number of observations per subject (k_i), where k_i was considered to be constant between subjects

$$\begin{bmatrix} n \\ k \end{bmatrix} = \begin{bmatrix} 100 & 300 & 600 \\ 10 & 50 & 100 \end{bmatrix}.$$

For correlations, we considered the variation of the within-subjects correlation (ρ) and the between-responses correlation (δ) at different levels: null (0%), low (5%), middle (50%) and high (95%)

$$\begin{bmatrix} \rho \\ \delta \end{bmatrix} = \begin{bmatrix} 0.00 & 0.05 & 0.50 & 0.95 \\ 0.00 & 0.05 & 0.50 & 0.95 \end{bmatrix}.$$

To study the effect of sample sizes or correlations on the performance of computational techniques, we randomly fixed a combination of correlations or a combination of sample size elements. The selected combinations were $\begin{bmatrix} n = 600 \\ k = 50 \end{bmatrix}$ and $\begin{bmatrix} \rho = 0.95 \\ \delta = 0.95 \end{bmatrix}$

for correlations and sample sizes studies, respectively. Simulations were conducted using R 3.5.0 (R Core Team, 2018) on a laptop Core i7 – 6500, CPU 2.59 GHz, RAM 12.00 GB. The function *mlmmm.em* of the *mlmmm* package (Schafer and Yucel, 2002) was used to perform FS-EM.

5. Comparison criteria of computational techniques

For each combination of sample size or correlation conditions, 200 longitudinal data sets were simulated and were used to compare computational techniques considered. We had 9 and 16 combinations respectively for sample size and correlation studies. In total, we simulated 4,800 data sets since the data set simulated with the characteristics $\begin{bmatrix} n = 600 & \rho = 0.95 \\ k = 50 & \delta = 0.95 \end{bmatrix}$

is common to both correlation and sample size studies. The comparison criteria (Table 1) were:

1. $\bar{\hat{\theta}}$, the mean of empirical estimates of parameters across the 200 simulations;
2. % Bias, the relative bias in estimating parameters;
3. RMSE, square root of mean square error of the estimator which stands for the measure of estimation accuracy as it takes into account both bias and variability (Zhou, 2009);
4. Reff, relative efficiency in estimating parameters;
5. Computing time (T);
6. N_r , the total number of iterations before reaching convergence (Zhou, 2009).

An Analysis of variance study is performed to assess effects of sample size and correlations variation on the estimation accuracy (RMSE) of methods, using the R package *stats* (R Core Team, 2018).

Table 1: Comparison statistics of computational techniques

N	Criteria statistics	Expressions
1	Mean	$\bar{\hat{\theta}} = \frac{\sum_{r=1}^{200} \hat{\theta}_r}{200}$
2	%Bias	$\%Bias = \frac{100 \times (\bar{\hat{\theta}} - \theta_T)}{ \theta_T }$
3	RMSE	$RMSE = \sqrt{\frac{\sum_{r=1}^{200} (\hat{\theta}_r - \theta_T)^2}{200}}$
4	Reff	$Reff_{CT_1, CT_2} = \frac{\sum_{r=1}^{200} (\hat{\theta}_{CT_1 r} - \theta_T)^2}{\sum_{r=1}^{200} (\hat{\theta}_{CT_2 r} - \theta_T)^2}$
5	Computing time	$T = \sum_{i=1}^{200} \frac{T_i}{200}$
6	Number of iterations	$N_r = \sum_{r=1}^{200} \frac{N_i}{200}$

$\hat{\theta}_r$, parameter estimated at the r^{th} simulation;
 θ_T , True value of a given parameter;
 CT , computational technique;
 T_i , number of iterations at convergence per simulation;
 N_i computing time per simulation.

6. Results

6.1. Effect of sample size on computational techniques' performance

The estimation accuracy of fixed effects, elements of random and residuals variance-covariance matrices varies significantly according to computational techniques and sample sizes. An exception was observed for the estimation of residuals' variance, which did not vary according to sample sizes (Table 2). For fixed effects estimation, CMLME and EM were the most accurate techniques, and FS-EM was the less accurate. Regarding residuals variance, EM and ECM methods provided the most accurate estimations while CMLME was the less accurate. The most accurate estimations for random effects were provided by ECM and CMLME while FS-EM was the less accurate technique. When the number of subjects increased (up to 600) the estimation accuracy of fixed and random effects decreased (Table 2).

The relative bias linked to fixed effects estimation was almost null for all computational techniques when the number of subjects was small (down to 100) irrespective to the number of observations per subject. From $n = 300$ to 600, fixed parameters were either underestimated or overestimated (Figure 1). Whatever the increase of sample size, CMLME and EM provided very slight relative bias (medians often null), while FS-EM often provided overestimations ($n = 300$) and underestimations ($n = 600$)(Figure 1). Concerning residuals variance and random effects estimations, ECM's relative bias was often null for any sample size (Figures 2 and 3). Almost null relative bias was recorded for all small sample sizes, while overestimations were observed for high number of subjects (300, 600). FS-EM and EM overestimated residuals variance and random effects (Figures 2 and 3).

Regarding fixed effects estimation, CMLME, ECM and FS-EM outperformed EM technique for all sample sizes except when the number of subjects was 300 coupled with 50 observations per subject (total sample size = 15000), in which case EM outperformed the other computational techniques. In general, the relative efficiency of computational techniques varied according to sample sizes and there is no best computational technique for all sample sizes. Though, CMLME outperformed EM, ECM and in particular FS-EM for most sample size situations (Figure 4). For the estimation of model residuals variance the four computational techniques considered provided similar relative efficiency. Although, EM technique outperformed FS-EM when numbers of observations per subject varied between 50 and 100. This difference of efficiency was particularly high, an order of 10^6 between both both techniques. In addition, for the smallest sample size considered in this study ($n = 100, k = 10$), FS-EM outperformed the other computational techniques. Meanwhile, for the highest sample size ($n = 600, k = 100$), ECM and CMLME outperformed FS-EM (Figure 4). Concerning random effects estimation, EM and FS-EM were the less efficient while ECM and CMLME were the most efficient for all sample size situations considered.

ECM and FS-EM recorded the most important computing time and the greatest number of iterations at convergence for any sample size (Figure 5). Detailed results concerning the effect of sample size on computational techniques' performance are available in appendix from Table 4 to 7.

Table 2: Least square means of RMSE according to sample size and computational methods.

	β			Σ			Γ		
	lsmean	F-statistics	P-value	lsmean	F-statistics	P-value	lsmean	F-statistics	P-value
Methods									
EM	2.61 ^b			2.42 ^b			43.51 ^b		
ECM	4.52 ^a	13.46	3.55e - 08***	3.39 ^b	4.99	0.003**	2.71 ^b	3.68	0.016*
FS-EM	5.21 ^a			68.27 ^b			139.73 ^a		
CMLME	2.36 ^b			231.28 ^a			3.07 ^b		
Sample size [n, k]									
[100, 10]	2.74 ^b			136.01 ^a			7.26 ^b		
[100, 50]	2.76 ^b			113.97 ^a			5.42 ^b		
[100, 100]	2.32 ^b			40.40 ^a			4.76 ^b		
[300, 10]	4.58 ^b			30.21 ^a			21.85 ^b		
[300, 50]	3.23 ^b	6.53	9.82e - 08***	71.33 ^a	0.56	0.809	11.67 ^b	4.73	0.000***
[300, 100]	4.32 ^b			24.63 ^a			7.43 ^b		
[600, 10]	7.04 ^a			18.13 ^a			3.74 ^b		
[600, 50]	3.36 ^b			82.16 ^a			23.70 ^b		
[600, 100]	2.74 ^b			170.24 ^a			339.43 ^a		

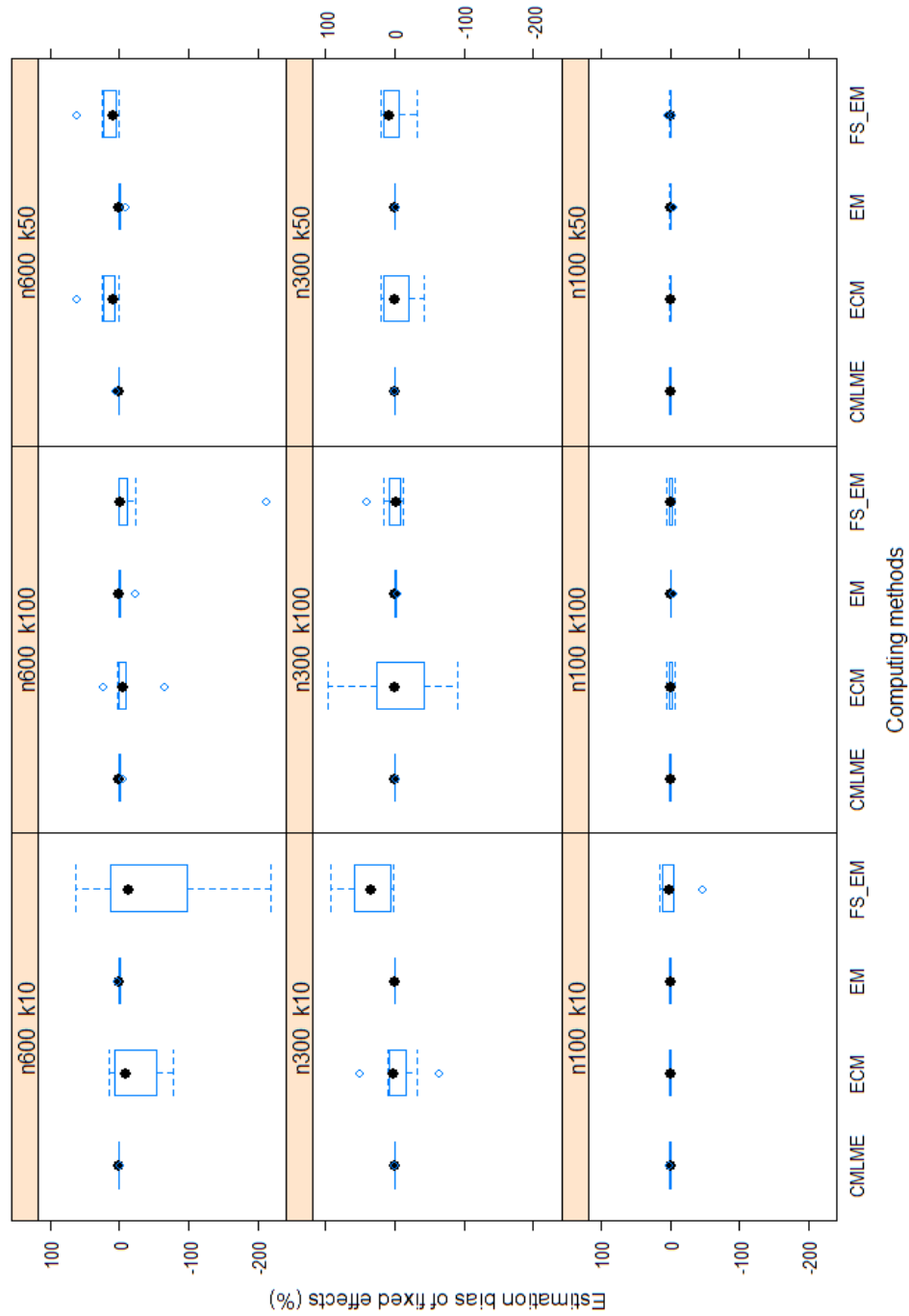


Fig. 1: Boxplot for the mean relative bias of fixed effects according to sample size variation and computational techniques.

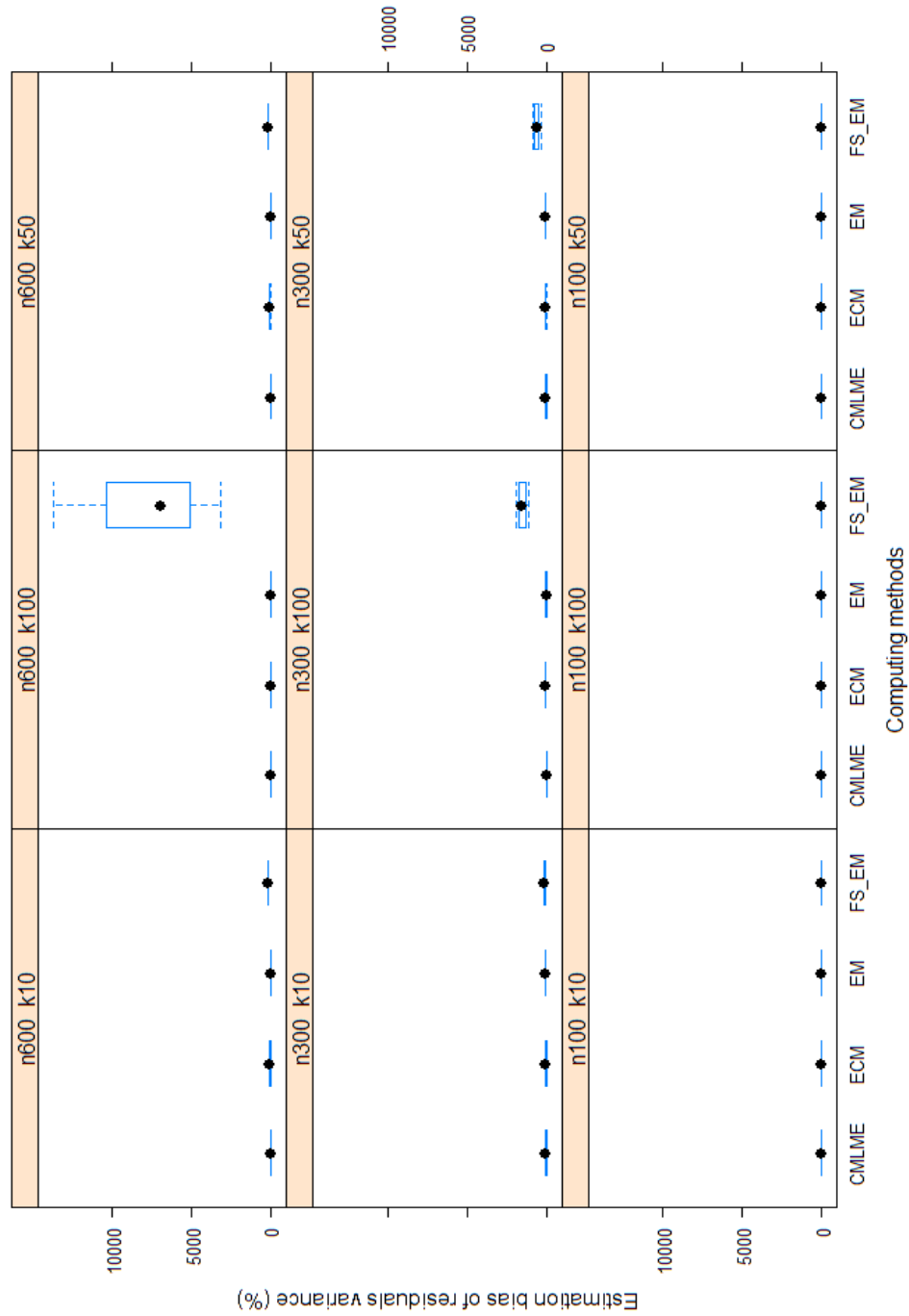


Fig. 2: Boxplot for the mean relative bias of residuals variance according to sample size variation and computational techniques.

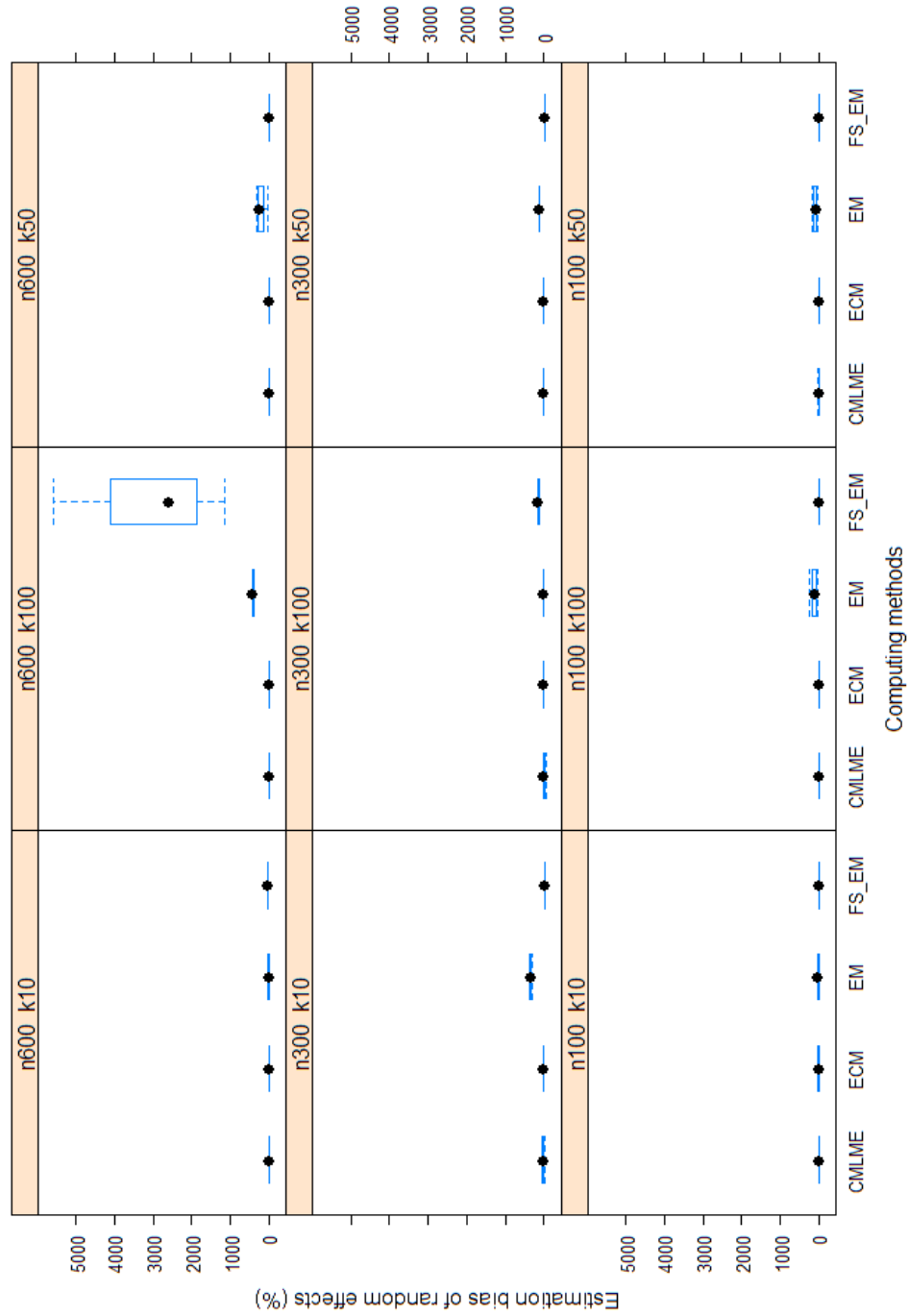


Fig. 3: Boxplot for the mean relative bias of random effects according to sample size variation and computational techniques.

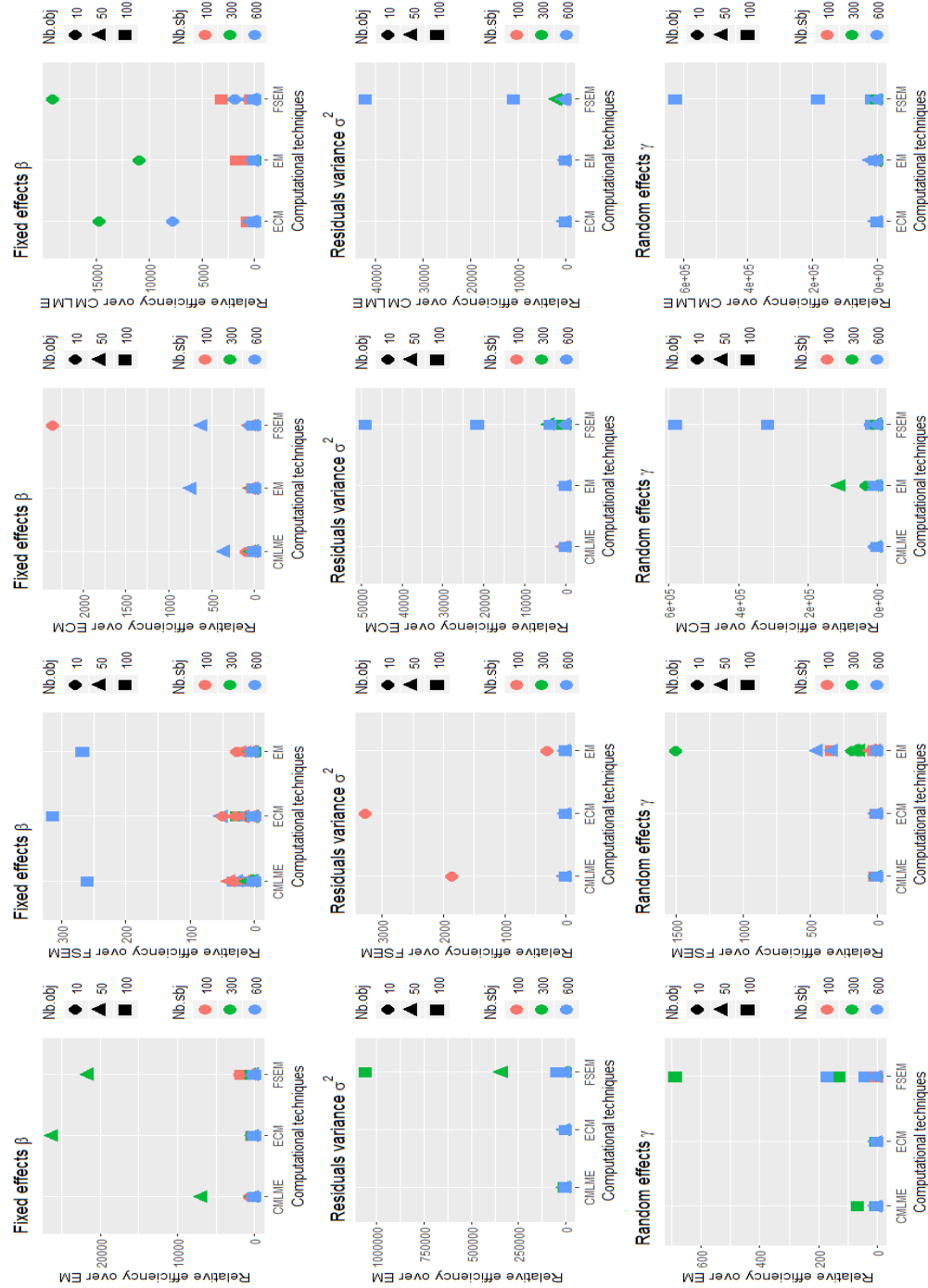


Fig. 4: Relative efficiency of computational techniques in sample size conditions. Each column of graphs represents the relative efficiency in estimating fixed effects, model's variance and random effects (up to down) of three computational techniques with respect to a given technique (EM, FS-EM, ECM and CMLME respectively, from left to right sides).

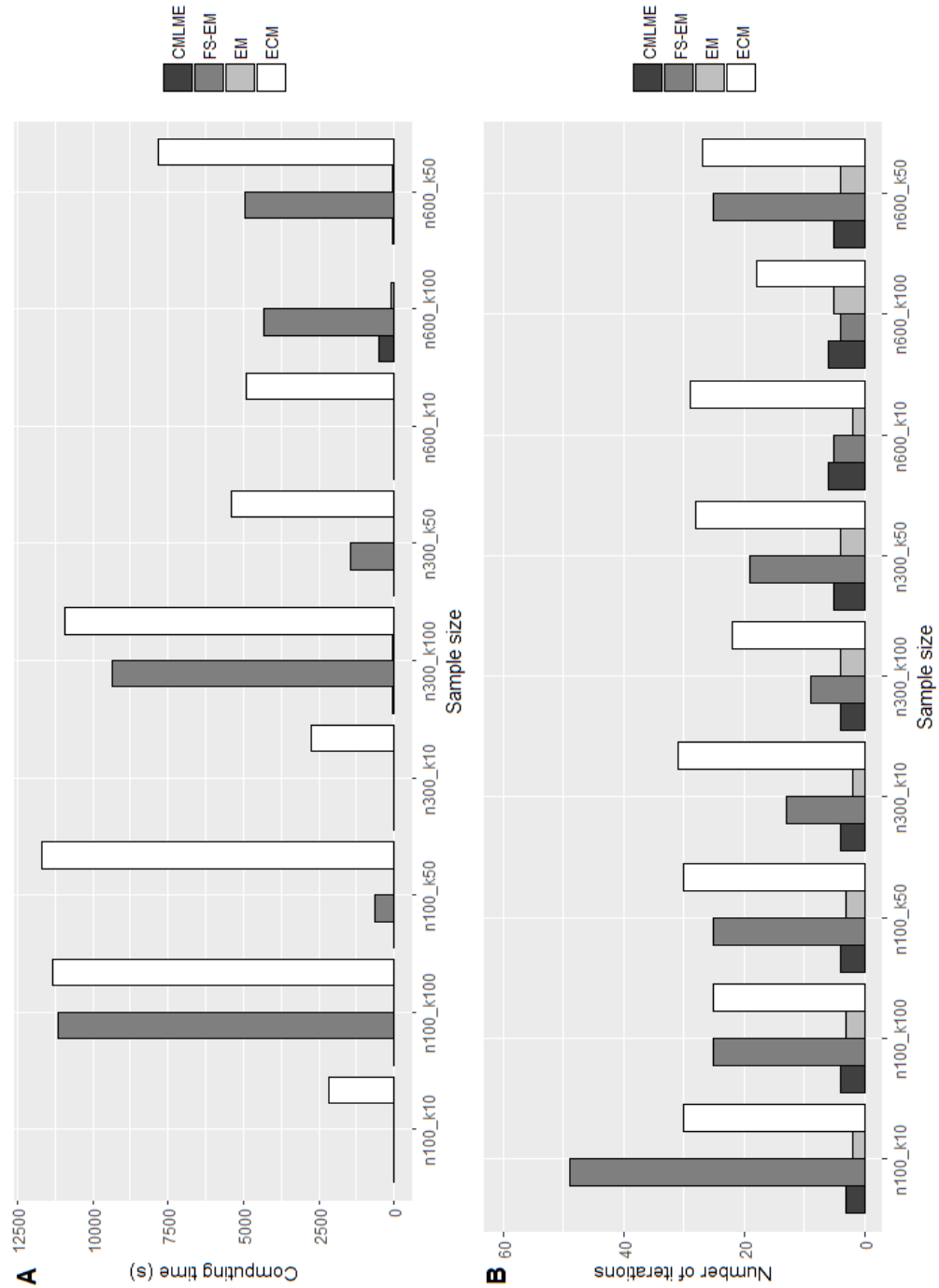


Fig. 5: **A** Average computational time variation according to sample size and computational techniques; **B** variation of average number of iterations at convergence according to sample size and computational techniques.

6.2. Effect of within-subjects and between-responses correlations on computational techniques' performance

The accuracy in estimating fixed and random effects and residuals variances varied significantly according to computational techniques and between-responses and within-subjects correlations (Table 3). CMLME and EM techniques provided the best accuracy for fixed effects while ECM and Fisher scoring were the most accurate for estimating random effects (Table 3). The lowest RMSE of residuals variance were recorded for EM and ECM and the highest for CMLME. In general, higher the within-subjects correlation ρ (0.95) was and better were estimations' accuracy of fixed and random effects. The contrary is observed for between-responses correlations (δ) for which the lowest average RMSE values were recorded for $\delta = 0.00, 0.05$ for any ρ value except for $\rho = 0.00$. All computational techniques provided best accuracy for all parameters when the between-responses correlation is low and the within-subjects correlation is high.

Concerning the relative bias in estimating fixed effects, only CMLME provided approximately null bias whatever the combination of correlations (Figure 6). ECM often underestimated fixed effects and provided sometimes overestimations with extreme bias especially for $[\rho = 0.00, \delta = 0.50]$ and $[\rho = 0.00, \delta = 0.95]$. ECM provided null relative bias for $[\rho = 0.05, \delta = 0.00]$ and $[\rho = 0.95, \delta = 0.95]$. EM provided null relative bias for $[\rho = 0.00, \delta = 0.50]$ and $[\rho = 0.05, \delta = 0.50]$. EM performed often better when the correlation between responses is around 0.50 than the other δ values. FS-EM generally underestimated fixed effects (Figure 6). Regarding residuals variance, the relative bias estimation did not vary according to correlations. Biases were in most cases null though FS-EM provided extreme biases (overestimations) (Figure 6). Concerning random factors, FS-EM and ECM were techniques with approximately null biases while EM often provided overestimations (Figure 7).

Variations of between-responses and within-subjects influenced the efficiency in estimating fixed effects, especially ECM and FS-EM, which provided less efficiency with regard to CMLME and EM techniques. All computational techniques provided similar efficiency in estimating model residuals variances. But FS-EM was relatively less efficient in $\rho = 0$ and $\delta = 5, 50$ conditions. Concerning random effects estimations, EM was the less efficient technique while ECM followed by CMLME were in general the most efficient techniques for all correlation conditions considered (Figure 9).

The computing time and number of iterations at convergence vary according to methods and correlations' variation (Figure 10). For all combinations of correlations, FS-EM method took much time to estimate parameters (1000 – 8500 seconds). In general, starting from low (0.05) to high (0.95) within-subject correlations, ECM method performed with the lowest rapidity (Figure 10A). CMLME performed very quickly for all correlation cases though it ran very low when both correlations were null (Figure 10A). Moreover, computing time is generally more considerable for combination cases which included either the null within-subjects correlation

($\rho = 0.00$) or the null between-responses correlation ($\delta = 0.00$) (Figure 10A). Regarding number of iterations at convergence, ECM and FS-EM methods recorded the highest values on contrary to EM and CMLME that recorded the lowest number of iterations (Figure 10B). Detailed results concerning the effect of correlations on computational techniques' performance are available in appendix from Table 8 to 11.

Table 3: Least square means of RMSE according to within-subject and between-responses correlations and computational methods.

Methods	β			Σ			Γ		
	lsmean	F-statistics	P-value	lsmean	F-statistics	P-value	lsmean	F-statistics	P-value
EM	2.34 ^b			2.79 ^b			34.46 ^a		
ECM	5.34 ^a	30.13	2.00e - 16 ^{***}	9.83 ^b	4.30	0.006 ^{**}	2.70 ^b	29.01	2.23e - 14 ^{***}
FS-EM	5.12 ^a			80.61 ^{ab}			3.84 ^b		
CMLME	2.62 ^b			134.52 ^a			4.31 ^b		
Correlations [ρ, δ]									
[0.00, 0.00]	4.06 ^{ab}			2.90 ^a			29.96 ^a		
[0.00, 0.05]	4.08 ^{ab}			204.14 ^a			12.67 ^{ab}		
[0.00, 0.50]	5.99 ^a			93.02 ^a			5.37 ^{ab}		
[0.00, 0.95]	3.54 ^{ab}			62.22 ^a			5.68 ^{ab}		
[0.05, 0.00]	3.08 ^b			3.12 ^a			7.30 ^{ab}		
[0.05, 0.05]	3.56 ^{ab}			36.84 ^a			9.53 ^{ab}		
[0.05, 0.50]	4.78 ^{ab}			24.92 ^a			6.97 ^{ab}		
[0.05, 0.95]	4.37 ^{ab}	1.91	0.021 [*]	138.91 ^a	0.87	0.593	7.55 ^{ab}	2.11	0.001 ^{**}
[0.50, 0.00]	4.03 ^{ab}			28.69 ^a			5.99 ^{ab}		
[0.50, 0.05]	3.59 ^{ab}			20.94 ^a			3.17 ^b		
[0.50, 0.50]	3.24 ^b			64.82 ^a			23.48 ^a		
[0.50, 0.95]	3.13 ^b			104.55 ^a			15.75 ^a		
[0.95, 0.00]	4.38 ^{ab}			10.40 ^a			3.72 ^b		
[0.95, 0.05]	2.57 ^b			4.07 ^a			3.40 ^b		
[0.95, 0.50]	3.92 ^{ab}			29.73 ^a			16.64 ^a		
[0.95, 0.95]	3.39 ^{ab}			81.76 ^a			24.01 ^a		

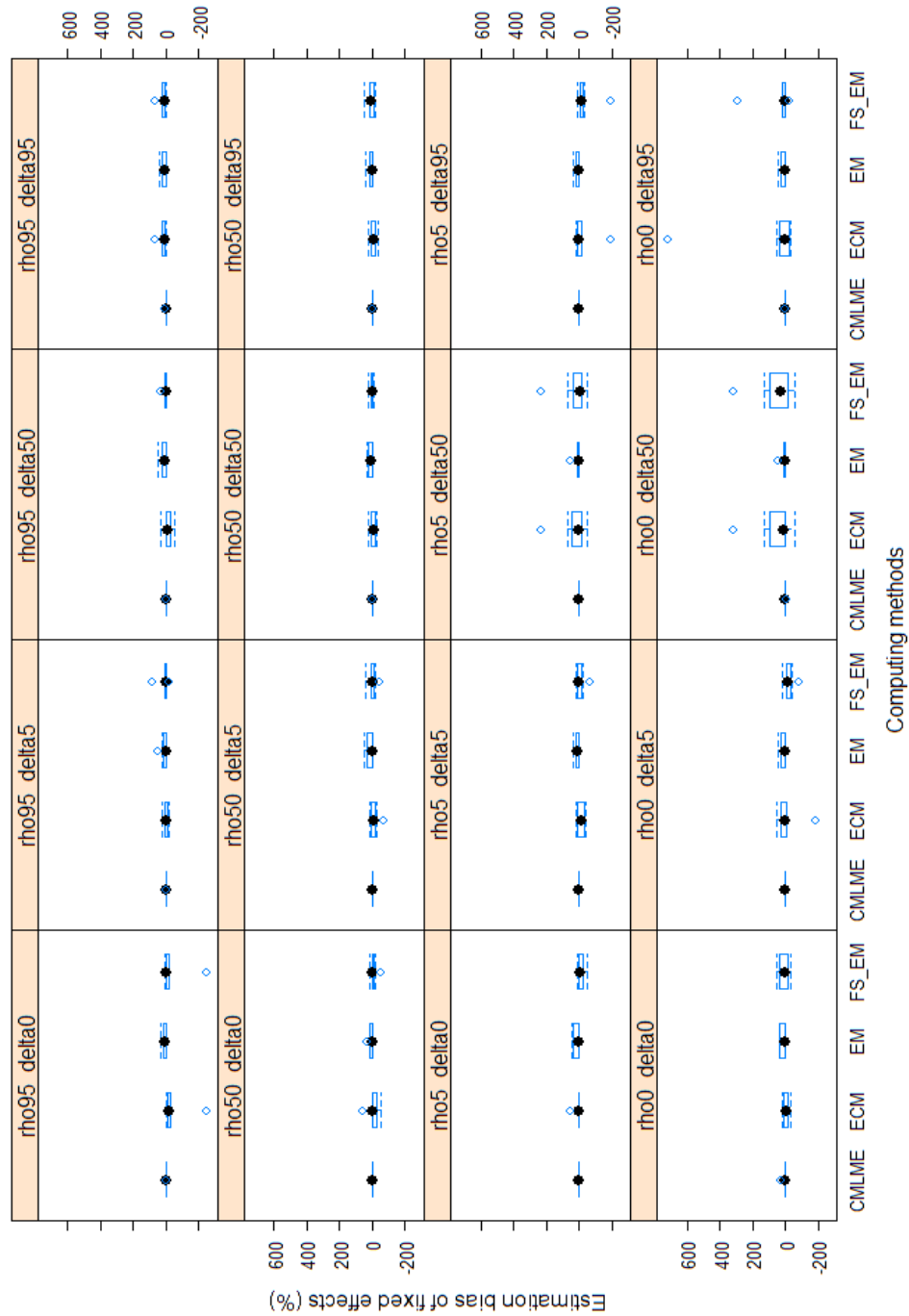


Fig. 6: Boxplot for the mean relative bias of fixed effects according to within and between correlations and computational techniques.

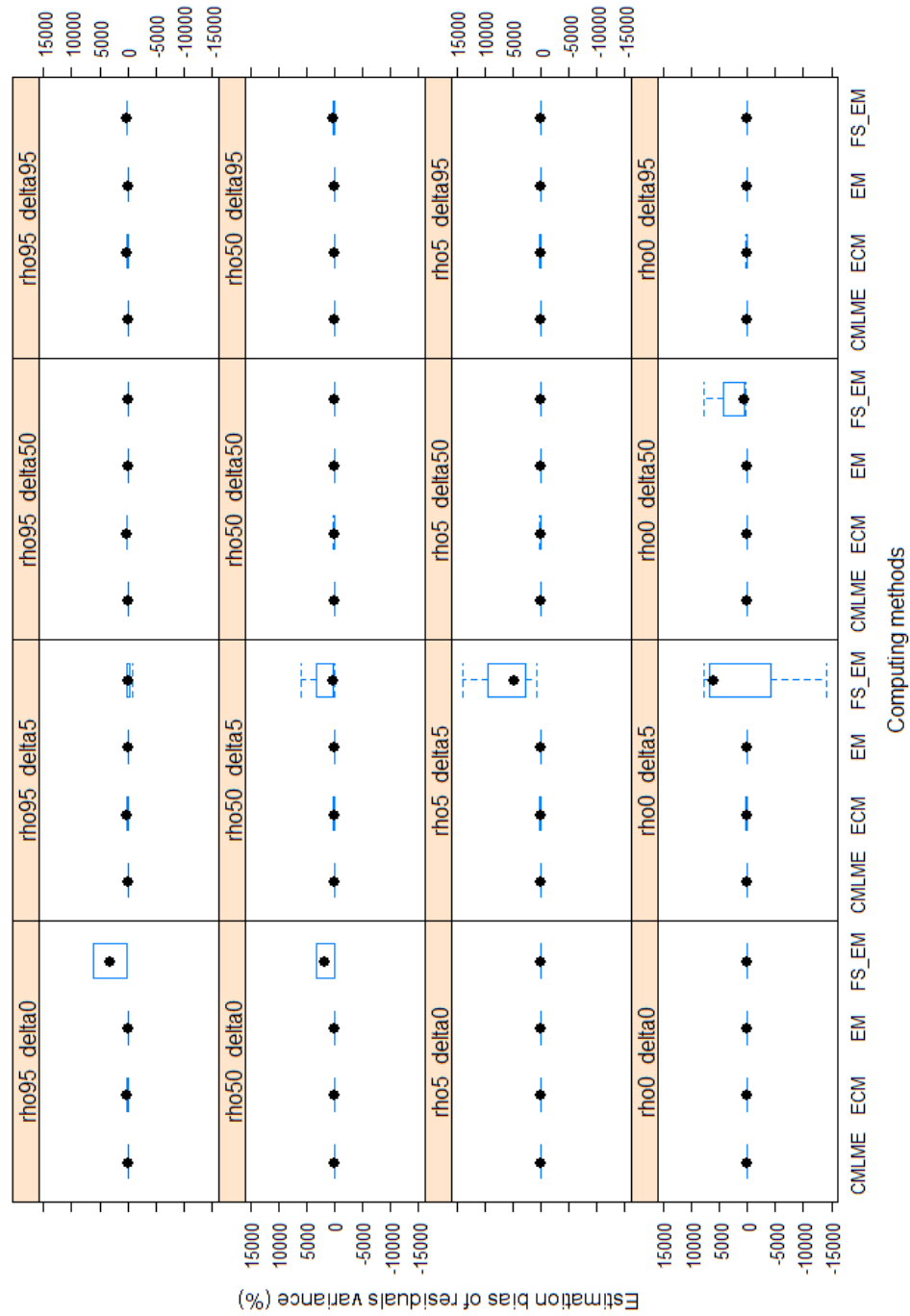


Fig. 7: Boxplot for the mean relative bias of residuals variance according to within and between correlations and computational techniques.

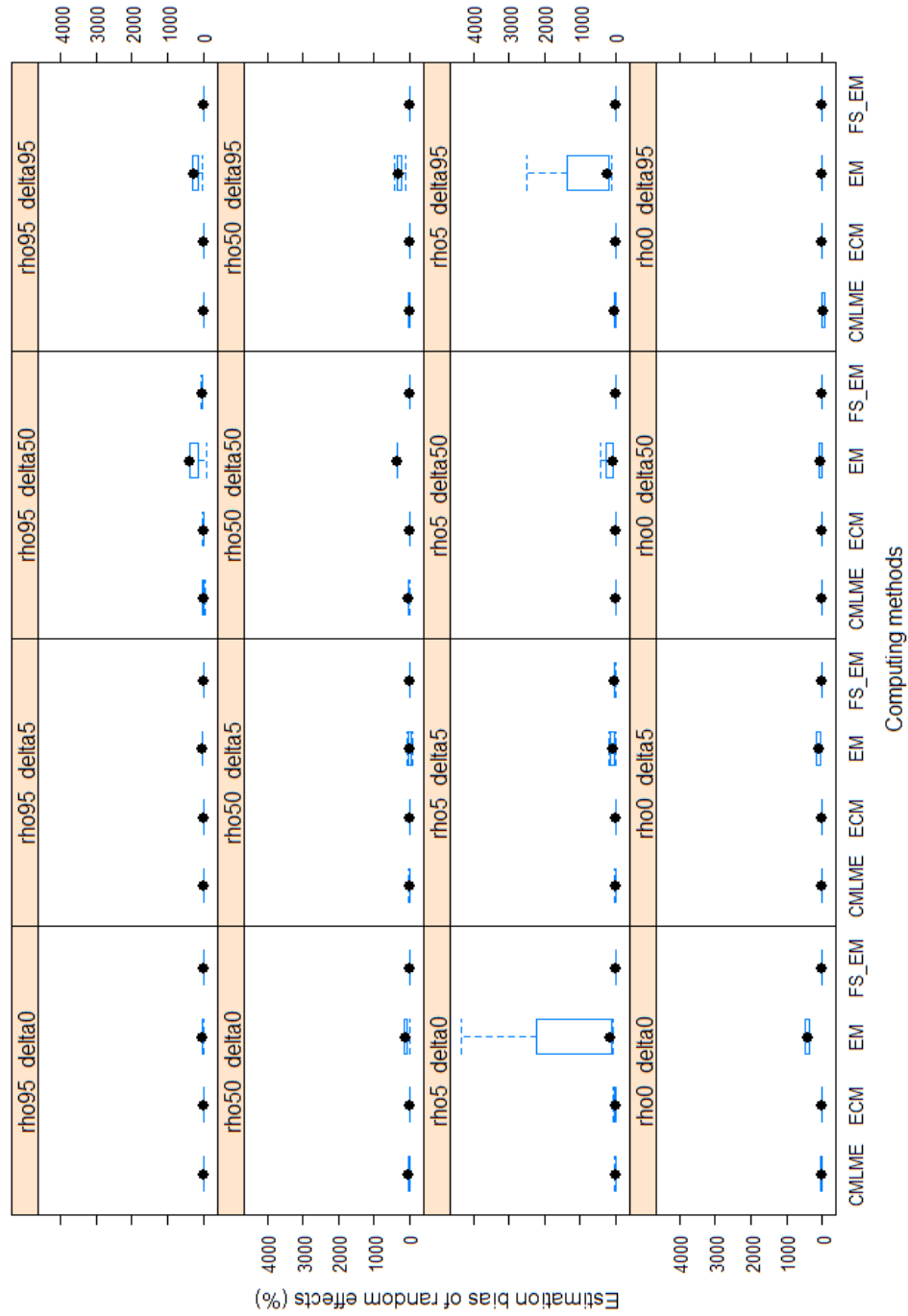


Fig. 8: Boxplot for the mean relative bias of random effects according to within and between correlations and computational techniques.

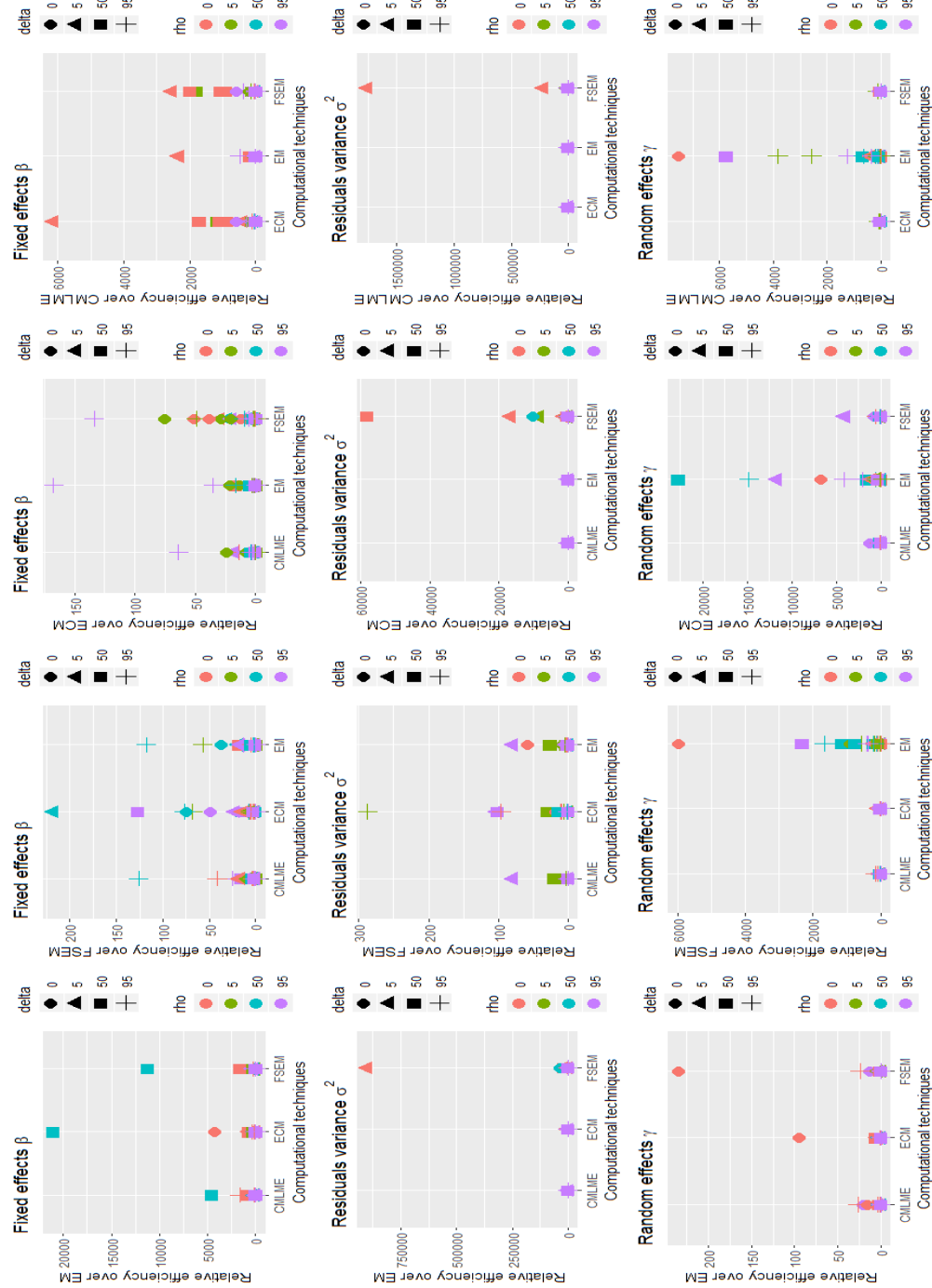


Fig. 9: Relative efficiency of computational techniques in correlation conditions. Each column of graphs represents the relative efficiency in estimating fixed effects, model's variance and random effects (up to down) of three computational techniques with respect to a given technique (EM, FS-EM, ECM and CMLME respectively, from left to right sides).

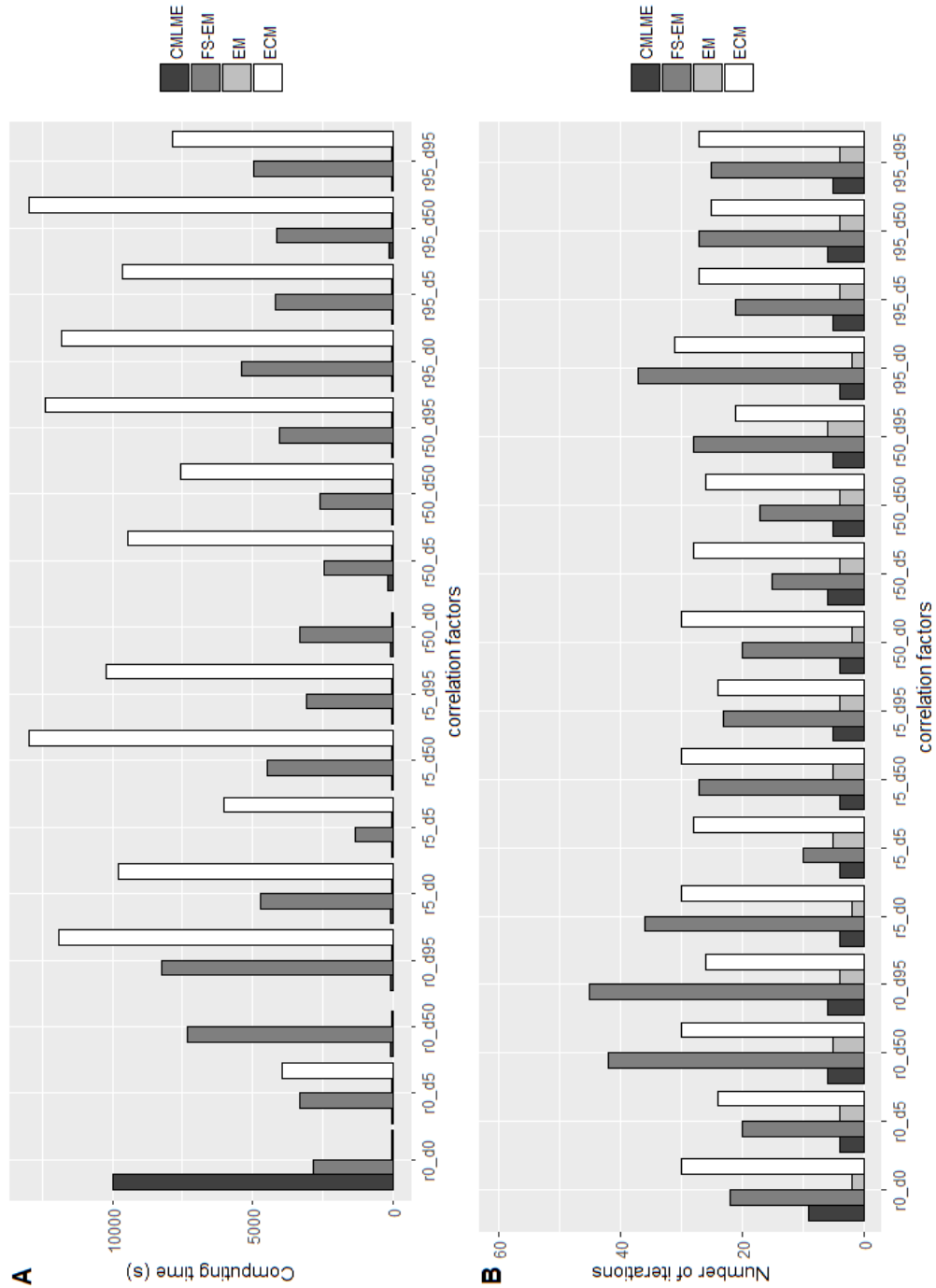


Fig. 10: **A** Average computational time variation according to within and between correlations and computational techniques; **B** variation of average number of iterations at convergence according to within and between correlations and computational techniques.

7. Discussion

7.1. Importance of computational techniques: common features and differences

In this study, we hypothesized that some methods can be more accurate than other regarding parameters to be estimated in MLMMs. Our findings showed that the best accuracy among computational techniques varied according to the nature of parameters to be estimated. Indeed, regarding fixed effects estimations, CMLME was the most accurate estimation method while ECM and EM were the most accurate for estimating variances' components, respectively. In general, CMLME and ECM were the best computational techniques though CMLME provided very important noises when the between-responses correlation is not null while ECM was less efficient in term of accuracy than CMLME. Adjakossa (2017) has shown that CMLME may be quite consistent in estimating both fixed and random effects' variances as this method uses Cholesky's factorization during random effects estimation procedure. But the comparison was done regarding traditional EM algorithm. Therefore, our findings revealed that these results from Adjakossa (2017) were actually valid but just for fixed effects, that can be due to the iterative updating of fixed effects vector. The fact that ECM outperforms CMLME in estimating random effects' variances can be due to the computation procedure that includes in each iteration an estimation of the auto-correlation matrix (Wang and Fan, 2010). Moreover, it takes simultaneously information concerning both variances between-responses and within-subjects. Thus, ECM estimates more efficiently random effects by taking into account more information regarding random variations than the other computational techniques. CMLME cannot estimate the covariance of the residuals variance-covariance matrix as currently the theoretical development is limited to a null correlation between responses. FS-EM was in general the less accurate method among the four computational techniques considered in the frame of this study. According to Schafer and Yucel (2002), FS-EM performs better than the traditional EM method. The difference between Schafer and Yucel (2002)'s results and our findings can be explained by the computational complications that FS-EM knows when the number of subjects becomes important. Indeed, we found out that FS-EM does not often reach convergence when at some scoring processes, the log-density function is not concave. We assume that without this drawback it can actually outperform the EM method. However, a non-negligible challenge was remarked and has to be taken into account for further improvements. In general, we consider predictors (fixed and random factors) to be constant for the whole MLMM and only responses may vary. But in practice, it is possible that covariates change across univariate LMM that build the MLMM. Indeed, on contrary to EM and CMLME, FS-EM and ECM do not allow a variation of covariates across univariate linear mixed effects models. A best computational method should combine the accuracy of CMLME, ECM and EM in estimating fixed effects, random effects and residuals variance covariance matrix, respectively.

7.2. Limits of MLMM computational techniques regarding variations of sample size and correlations

We assessed variations of parameters' estimation accuracy according to sample sizes. Results showed a dependence of estimations' accuracy and computation time to the variation of the sample size especially the number of subjects. The difference of accuracy and computing time appears striking either from $n = 300$ or when the total sample size (N) exceeds 10000. Regarding correlations, there is also a variation of the accuracy according to the combination of the between-responses (δ and within-subjects (ρ) correlations. Most of the computational techniques' accuracy was low when both ρ and δ were null. Wang and Fan (2010) also showed that ECM algorithm for fitting MLMM performs better when the sample size is small and the between-responses correlation is large. More, other studies (Mingers, 1989; Tekwe *et al.*, 2004) confirmed that the variation of computational factors have significant effect on parameters' estimations. As in general, accuracy of computational techniques in estimating MLMMs parameters decreases as well as sample size increases, and regarding the highest sample size considered in this study (60000 observations per variable to be considered in the model), computational techniques may be less accurate and efficient in presence of big data. The challenge is therefore to adjust existing techniques in order to improve accuracy when data become very big.

8. Conclusion

This simulation study tackled the performance of four (two old and two recent) major computational techniques used for fitting MLMMs regarding variations of some data features (sample size and correlations). We assumed there exist significant effects of these factors on their estimations' accuracy and computing time. In general, larger the between-responses and within-subjects correlations are, better computational techniques performed. Big sample size introduces less accuracy of methods in estimating parameters though CMLME is in these cases the most reliable. Our initial assumption was thus confirmed and CMLME and ECM were pointed out as the most accurate computational techniques. Though, CMLME is the most accurate for fixed effects estimations while ECM is the most accurate for random effects estimations. Therefore, there is not an absolute best computational method among the four we considered in the frame of this study. But as ECM often performs very slowly especially when the sample size increased, we can rather consider CMLME as the core method that may be improved using some features of the other techniques. Consequently, an extension of the CMLME that includes residuals heteroscedasticity is quite required for improving its estimation accuracy. The best computational technique may also be accurate and efficient for big data whatever the between-responses and within-subjects correlations considered.

Acknowledgments

Authors acknowledge the the Agropolis Foundation, Fondazione Cariplo and Fondation Daniel & Nina Carasso, for funding the project TREEFOOD n°1507 – 143 in which this study was performed.

References

- Adjakossa, E.H., Sadissou, I., Hounkonnou, M.N., and Nuel, G., 2016. Multivariate Longitudinal Analysis with Bivariate Correlation Test. *PLoS one*, 11(8), e0159649. <https://doi.org/10.1371/journal.pone.0159649>.
- Adjakossa, E.H., 2017. Analyse longitudinale multivariée par modèles mixtes et application à l'épidémie de la malaria. *PhD thesis*, Paris 6.
- An, X., Yang, Q., and Bentler, P.M., 2013. A latent factor linear mixed model for high-dimensional longitudinal data analysis. *Statistics in Medicine*, doi: 10.1002/sim.5825.
- Barndorff-Nielsen, O., and Schou, G., 1973. On the parametrization of autoregressive models by partial autocorrelations. *Journal of multivariate Analysis*, 3(4), 408-419.
- Box, G.E.P., Jenkins, G.M., and Reinsel, G.C., 1994. *Time series analysis: forecasting and control*. Prentice Hall.
- Box, G.E.P., Jenkins, G.M., Reinsel, G.C., and Ljung, G.M., 2015. *Time series analysis: forecasting and control*. John Wiley & Sons.
- Chi, E.M., and Reinsel, G.C., 1989. Models for longitudinal data with random effects and AR (1) errors. *Journal of the American Statistical Association*, 84(406), 452-459.
- Dempster, A.P., Laird, N.M., and Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society. Series B (methodological)*, 1-38.
- Fieuws, S., and Verbeke, G., 2006. Pairwise fitting of mixed models for the joint modeling of multivariate longitudinal profiles. *Biometrics*, 62(2), 424-431.
- Fieuws, S., Verbeke, G., and Molenberghs, G., 2007. Random-effects models for multivariate repeated measures. *Biometrics*, 16(5), 387-397.
- Jennrich, R.I., and Schluchter, M.D., 1986. Unbalanced repeated-measures models with structured covariance matrices. *Biometrics*, 805-820.
- Laird, N.M., and Ware, J.H., 1982. Random-effects models for longitudinal data. *Biometrics*, 963-974.
- Laird, N.M., Lange, N., and Stram, D., 1987. Maximum likelihood computations with repeated measures: application of the EM algorithm. *Journal of the American Statistical Association*, 82(397), 97-105.
- Lindstrom, M.J., and Bates, D.M., 1988. Newton—Raphson and EM algorithms for linear mixed-effects models for repeated-measures data. *Journal of the American Statistical Association*, 83(404), 1014-1022.
- Meng, X-L., and Rubin, D.B., 1988. Maximum likelihood estimation via the ECM algorithm: A general framework. *Biometrika*, 80(2), 267-278.
- Mingers, J., 1989. Maximum likelihood estimation via the ECM algorithm: A general framework. *Biometrika*, Vienna, Austria, 227-243.

- R Core Team, 2018. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*, 4(2), <https://www.R-project.org/>.
- Reeves, J., and MacKenzie, G., 1998. A bivariate regression model with serial correlation. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47(4), 607-615.
- Reinsel, G., 1982. Multivariate repeated-measurement or growth curve models with multivariate random-effects covariance structure. *Journal of the American Statistical Association*, 77(377), 190-195.
- Reinsel, G., 1984. Estimation and prediction in a multivariate random effects generalized linear model. *Journal of the American Statistical Association*, 79(386), 406-414.
- Roy, J. and Lin, X., 2000. Latent variable models for longitudinal data with multiple continuous outcomes. *Biometrics*, 56(4), 1047-1054.
- Sammel, M., Lin, X., and Ryan, L., 1999. Multivariate linear mixed models for multiple outcomes. *Statistics in medicine*, 18(1718), 2479-2492.
- Schafer, J.L., and Yucel, R.M., 2002. Computational strategies for multivariate linear mixed-effects models with missing values. *Journal of computational and Graphical Statistics*, 11(2), 437-457.
- Shah, A., Laird, N., and Schoenfeld, D., 1997. A random-effects model for multiple characteristics with possibly missing data. *Journal of the American Statistical Association*, 92(438), 775-779.
- Tekwe, C.D., Carter, R.L., Ma, C-X., Algina, J., Lucas, M.E., Roth, J., Ariet, M., Fisher, T., and Resnick, M.B., 2004. An empirical comparison of statistical models for value-added assessment of school performance. *Journal of Educational and Behavioral Statistics*, 29(1), 11-36.
- Verbeke, G., Fieuws, S., Molenberghs, G., and Davidian, M., 2014. The analysis of multivariate longitudinal data: A review. *Statistical methods in medical research*, 23, 42-59.
- Wang, W-L., and Fan, T-H., 2010. ECM-based maximum likelihood inference for multivariate linear mixed models with autoregressive errors. *Computational Statistics & Data Analysis*, 54(5), 1328-1341.
- Zhou, M., 2009. Fully exponential Laplace approximation EM algorithm for non-linear mixed effects models. *PhD thesis*, University of Nebraska-Lincoln.

9. Appendix

Table 4: Empirical mean, percentage of bias and square root of mean square error of the fixed effects' estimators regarding the first response variable according to sample size.

Parameters	n	k	θ_T	EM			Fisher scoring - EM			ECM			CMLME		
				$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE
$\beta_1^{(1)}$	100	10	48.66	48.42	-0.50	4.90	48.49	-0.35	0.95	48.49	-0.35	2.50	48.65	-0.03	0.59
	100	50	17.67	17.58	-0.50	3.86	17.53	-0.78	1.10	17.53	-0.79	4.22	17.66	-0.03	1.73
	100	100	43.08	42.34	-1.71	1.67	42.73	-0.81	4.08	42.73	-0.81	4.12	43.05	-0.08	1.09
	300	10	36.31	36.08	-0.63	3.74	36.74	1.18	4.92	35.97	-0.94	1.55	36.30	-0.03	0.09
	300	50	18.19	18.09	-0.54	3.30	19.37	6.46	2.63	17.02	-6.44	3.57	18.18	-0.08	1.74
	300	100	35.81	35.56	-0.71	3.62	31.36	-12.44	6.00	38.13	6.47	2.57	35.78	-0.09	0.54
	600	10	39.48	39.15	-0.84	1.91	35.84	-9.22	4.98	41.79	5.86	2.76	39.47	-0.03	4.63
	600	50	45.15	44.43	-1.61	1.60	54.37	20.42	10.47	54.37	20.42	9.64	45.33	0.40	0.65
	600	100	35.49	34.89	-1.68	1.75	27.36	-22.92	9.40	34.45	-2.94	2.78	35.35	-0.41	3.59
$\beta_2^{(1)}$	100	10	42.79	42.58	-0.50	1.01	40.77	-4.72	3.02	42.67	-0.28	1.19	42.69	-0.23	2.59
	100	50	20.41	20.35	-0.28	1.50	19.80	-2.99	4.54	20.40	-0.05	1.99	20.44	0.15	4.69
	100	100	15.33	15.31	-0.13	2.49	15.13	-1.30	3.90	15.10	-1.50	1.79	15.34	0.07	0.07
	300	10	10.03	10.09	0.60	1.54	19.38	93.22	9.57	15.16	51.15	5.84	10.03	0.00	2.92
	300	50	19.79	19.84	0.25	2.07	23.81	20.31	4.62	23.78	20.16	4.75	19.74	-0.23	4.99
	300	100	4.27	4.26	-0.23	3.71	4.92	15.22	1.44	8.36	95.78	5.98	4.28	0.25	2.46
	600	10	32.29	32.30	0.03	1.78	43.79	35.61	11.78	8.35	-74.14	23.78	32.30	0.03	0.27
	600	50	5.28	5.30	0.38	0.61	5.57	5.50	0.32	5.75	8.90	2.26	5.30	0.47	1.62
	600	100	27.16	27.27	0.41	1.22	27.24	0.29	1.07	28.04	3.24	4.76	27.16	0.00	2.75
$\beta_3^{(1)}$	100	10	10.37	10.33	-0.39	2.23	11.19	7.91	2.84	10.38	0.10	2.18	10.37	0.00	0.60
	100	50	35.35	35.40	0.14	0.18	35.25	-0.28	3.63	35.40	0.14	2.25	35.39	0.11	2.45
	100	100	15.61	15.61	0.00	0.08	14.47	-7.30	3.33	14.56	-6.73	1.33	15.63	0.13	0.38
	300	10	21.66	21.55	-0.50	4.10	23.28	7.48	2.13	14.56	-32.78	7.11	21.65	-0.05	0.44
	300	50	4.04	4.04	0.09	5.17	3.75	-7.10	2.19	2.35	-41.87	1.75	4.04	0.00	3.76
	300	100	27.24	27.24	0.00	4.39	27.14	-0.37	4.76	2.50	-90.84	25.25	27.19	-0.18	3.45
	600	10	34.50	34.49	-0.03	5.20	30.40	-11.88	6.36	26.76	-22.43	7.82	34.51	0.03	3.40
	600	50	34.98	35.03	0.14	2.87	38.34	9.61	4.99	38.35	9.63	3.36	35.02	0.11	1.79
	600	100	45.73	45.82	0.19	4.65	45.68	-0.11	4.20	44.80	-2.03	1.53	45.73	0.00	2.17
$\beta_4^{(1)}$	100	10	18.94	19.21	1.43	3.28	19.57	3.33	0.63	19.11	0.90	4.44	19.08	0.74	3.20
	100	50	31.33	31.26	-0.21	2.94	32.01	2.17	3.92	31.30	-0.10	4.47	31.43	0.32	3.55
	100	100	34.07	34.02	-0.16	1.49	35.72	4.84	2.92	35.66	4.67	1.64	34.03	-0.12	0.86
	300	10	32.37	32.44	0.22	0.64	50.82	57.00	18.84	35.69	10.26	3.40	32.40	0.09	0.92
	300	50	16.07	16.12	0.31	2.64	19.15	19.17	3.68	19.14	19.10	3.77	16.15	0.50	3.32
	300	100	12.17	12.11	-0.49	4.68	10.77	-11.50	3.62	5.12	-57.93	7.40	12.17	0.00	3.44
	600	10	22.74	22.75	0.04	1.27	2.31	-89.84	20.87	5.08	-77.66	17.62	22.74	0.00	0.85
	600	50	27.08	27.16	0.29	2.77	33.89	25.15	7.39	33.89	25.15	7.43	27.09	0.04	3.12
	600	100	13.62	13.58	-0.32	1.65	13.52	-0.73	0.68	12.48	-8.37	1.35	13.63	0.07	3.87

Table 5: Empirical mean, percentage of bias and square root of mean square error of the fixed effects' estimators regarding the second response variable according to sample size.

Parameters	n	k	θ_T	EM			Fisher scoring - EM			ECM			CMLME		
				$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE
$\beta_1^{(2)}$	100	10	24.83	24.57	-1.05	1.99	24.04	-3.20	0.82	24.77	-0.23	4.49	24.82	-0.06	4.60
	100	50	5.24	5.03	-3.96	4.90	5.21	-0.55	4.52	5.21	-0.58	1.39	5.23	-0.20	2.18
	100	100	4.84	4.57	-5.50	3.41	4.66	-3.77	5.06	4.66	-3.78	3.44	4.82	-0.34	4.10
	300	10	48.23	47.88	-0.73	3.93	55.10	14.24	7.48	48.00	-0.47	1.42	48.22	-0.02	3.76
	300	50	5.87	5.69	-3.04	1.70	3.96	-32.61	3.23	3.96	-32.60	5.47	5.87	0.02	2.42
	300	100	5.35	5.09	-4.94	4.23	7.52	40.56	3.16	7.52	40.61	4.36	5.32	-0.50	2.52
	600	10	39.00	38.81	-0.49	2.32	32.94	-15.55	6.46	41.14	5.49	4.60	38.98	-0.05	3.52
	600	50	5.06	4.60	-9.16	1.71	8.17	61.53	3.10	8.17	61.53	5.37	5.28	4.35	2.66
	600	100	1.21	0.94	-22.58	5.10	-1.37	-213.13	2.89	0.43	-64.67	1.82	1.15	-4.65	1.43
$\beta_2^{(2)}$	100	10	22.20	22.04	-0.72	0.72	11.51	-48.15	11.14	22.10	-0.45	0.23	22.14	-0.27	2.17
	100	50	18.18	18.19	0.06	3.16	18.12	-0.33	3.48	18.12	-0.33	0.52	18.13	-0.28	3.16
	100	100	10.23	10.23	0.00	3.46	10.19	-0.39	2.10	10.17	-0.59	1.66	10.26	0.29	1.30
	300	10	29.05	29.11	0.21	2.01	44.96	54.77	15.91	10.26	-64.68	18.75	29.08	0.10	2.70
	300	50	48.37	48.40	0.06	0.04	53.82	11.27	5.86	54.74	13.16	6.46	48.40	0.07	3.28
	300	100	28.79	28.78	-0.03	4.00	29.46	2.33	4.58	31.97	11.06	6.04	28.79	0.00	4.68
	600	10	30.49	30.70	0.68	1.68	49.70	63.00	19.67	32.61	6.95	4.50	30.47	-0.07	0.46
	600	50	37.28	37.25	-0.08	4.88	37.44	0.43	4.44	37.44	0.43	0.18	37.34	0.16	3.32
	600	100	10.35	10.34	-0.05	4.25	10.53	1.74	0.26	12.68	22.51	4.60	10.30	-0.48	4.19
$\beta_3^{(2)}$	100	10	28.54	28.47	-0.25	0.12	32.85	15.10	5.20	28.50	-0.14	0.67	28.53	-0.02	3.02
	100	50	20.47	20.46	-0.04	4.14	20.43	-0.20	0.93	20.35	-0.61	1.67	20.46	-0.05	3.22
	100	100	49.34	49.34	0.00	2.57	48.92	-0.85	2.58	49.72	0.78	5.08	49.41	0.14	4.62
	300	10	47.11	47.11	0.00	4.17	49.82	5.75	5.52	49.06	4.14	4.84	47.13	0.04	0.04
	300	50	8.62	8.62	0.02	2.38	8.13	-5.69	1.18	8.06	-6.50	0.82	8.63	0.12	3.60
	300	100	4.86	4.74	-2.47	1.51	4.76	-2.06	3.17	4.46	-8.23	2.03	4.78	-1.65	1.92
	600	10	6.50	6.63	2.00	5.34	-0.28	-104.31	7.94	4.44	-31.69	4.97	6.50	0.00	2.76
	600	50	13.88	13.84	-0.29	2.63	14.93	7.56	3.07	14.97	7.85	1.72	13.87	-0.07	2.76
	600	100	39.26	39.23	-0.08	0.99	39.13	-0.33	0.91	36.69	-6.55	3.29	39.31	0.13	1.53
$\beta_4^{(2)}$	100	10	23.44	23.80	1.54	4.11	26.63	13.61	4.87	23.79	1.49	5.22	23.75	1.32	2.10
	100	50	27.39	27.51	0.44	1.15	27.43	0.15	0.60	27.31	-0.29	2.54	27.52	0.46	3.68
	100	100	15.21	15.17	-0.26	2.04	15.81	3.94	0.93	15.64	2.83	0.51	15.25	0.26	0.05
	300	10	5.37	5.35	-0.37	0.60	8.53	58.78	4.80	5.73	6.62	2.48	5.39	0.37	0.46
	300	50	49.44	49.62	0.36	0.35	53.57	8.35	4.44	53.53	8.27	4.43	49.54	0.20	3.79
	300	100	24.71	24.69	-0.08	0.57	23.52	-4.82	3.74	17.91	-27.52	7.29	24.71	0.00	1.00
	600	10	15.55	15.47	-0.51	4.29	-18.51	-219.04	34.03	17.99	15.69	4.73	15.48	-0.45	2.75
	600	50	34.86	34.71	-0.43	1.56	36.97	6.05	3.43	36.96	6.02	4.67	34.87	0.03	0.98
	600	100	33.13	33.14	0.03	3.28	32.94	-0.57	1.88	30.39	-8.27	3.25	33.12	-0.03	0.58

Table 6: Empirical mean, percentage of bias and square root of mean square error of the estimators of residuals' matrix of covariance according to sample size.

Parameters	n	k	θ_r	EM			Fisher scoring - EM			ECM			CMLME		
				$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE
σ_1^2	100	10	17.89	18.18	1.63	1.40	17.97	0.43	0.08	17.89	0.00	4.57	17.87	-0.09	3.44
	100	50	18.92	19.03	0.57	0.51	19.01	0.45	1.50	18.92	0.00	0.19	18.31	-3.25	3.91
	100	100	3.76	3.78	0.47	2.65	5.36	42.65	2.69	3.77	0.38	2.11	3.89	3.46	4.12
	300	10	3.61	3.81	5.67	2.25	5.38	49.05	4.14	3.62	0.40	3.66	3.50	-3.13	3.98
	300	50	20.25	20.44	0.96	3.64	136.69	575.00	116.49	20.31	0.31	1.88	20.17	-0.38	2.81
	300	100	2.89	2.89	0.02	0.03	33.72	1066.82	30.83	3.65	26.33	1.09	2.89	-0.05	2.61
	600	10	1.25	1.68	34.78	0.84	3.18	154.59	4.22	2.39	91.11	3.37	1.32	5.27	2.86
	600	50	19.27	19.67	2.07	4.99	56.13	191.30	36.91	30.74	59.54	11.81	19.26	-0.04	1.78
σ_{12}^2	600	100	4.41	4.43	0.48	4.00	145.46	3198.47	141.03	4.57	3.60	2.28	4.41	-0.02	1.35
	100	10	15.99	16.25	1.62	4.60	16.43	2.78	3.02	15.99	0.00	3.45	-	-	-
	100	50	13.39	13.49	0.72	5.07	13.40	0.06	3.96	13.39	0.00	4.77	-	-	-
	100	100	4.53	4.55	0.52	1.17	5.16	13.89	3.14	4.57	0.94	3.32	-	-	-
	300	10	3.30	3.67	11.18	1.34	6.33	91.70	3.43	3.34	1.29	0.53	-	-	-
	300	50	6.75	6.78	0.51	0.04	29.65	339.24	23.08	7.60	12.61	2.47	-	-	-
	300	100	1.92	1.93	0.42	0.66	30.93	1510.77	29.19	2.64	37.72	2.76	-	-	-
	600	10	1.84	2.02	9.57	2.28	5.07	175.57	5.92	2.56	39.38	1.26	-	-	-
σ_2^2	600	50	8.88	8.93	0.54	3.45	24.61	177.11	15.67	15.07	69.76	8.06	-	-	-
	600	100	5.31	5.33	0.39	5.24	375.64	6974.17	370.12	5.67	6.82	2.52	-	-	-
	100	10	15.84	16.22	2.39	4.40	16.49	4.08	2.16	15.84	0.00	4.55	15.77	-0.41	1.65
	100	50	10.50	10.65	1.47	1.24	10.50	0.00	2.65	10.50	0.00	3.40	10.65	1.40	1.56
	100	100	6.05	6.08	0.43	4.19	7.04	16.36	2.26	6.06	0.24	0.95	7.30	20.72	5.24
	300	10	3.35	4.17	24.56	0.96	8.53	154.74	5.24	3.36	0.43	4.19	4.20	25.43	2.90
	300	50	2.50	2.60	3.80	0.19	23.97	858.78	21.71	3.65	45.84	3.36	3.22	28.82	5.36
	300	100	1.42	1.44	1.11	2.79	28.84	1930.86	27.42	2.11	48.25	4.61	1.00	-29.48	1.58
600	10	2.99	3.12	4.28	1.93	8.38	180.43	6.32	3.67	22.83	1.17	2.98	-0.25	3.32	
600	50	4.54	4.60	1.43	1.02	12.75	180.83	8.39	5.67	24.93	4.82	3.59	-20.82	1.15	
600	100	7.08	7.14	0.90	4.50	978.76	13724.28	971.69	8.04	13.58	4.40	5.23	-26.19	4.74	

Table 7: Empirical mean, percentage of bias and square root of mean square error of the estimators of random effects' matrix of covariance according to sample size.

Parameters	n	k	θ_T	EM			Fisher scoring - EM			ECM			CMLME		
				$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE
Γ_1	100	10	79.21	90.29	13.99	11.14	71.31	-9.98	8.67	83.75	5.73	5.02	79.16	-0.06	4.37
	100	50	8.52	24.60	188.77	16.17	7.68	-9.91	2.40	8.97	5.32	0.67	10.41	22.14	2.15
	100	100	4.57	16.36	258.05	12.30	4.11	-10.12	0.67	4.70	2.92	1.63	4.10	-10.24	2.50
	300	10	15.61	64.11	310.69	48.67	14.04	-10.08	4.35	15.73	0.79	3.12	13.65	-12.56	2.14
	300	50	23.34	52.34	124.27	29.17	21.02	-9.95	2.60	23.36	0.09	0.56	23.30	-0.19	1.58
	300	100	18.37	22.40	21.92	4.72	40.75	121.84	22.50	18.44	0.36	3.59	18.31	-0.35	0.54
	600	10	14.56	20.27	39.21	7.15	17.80	22.24	3.24	14.63	0.48	4.24	15.21	4.50	1.32
	600	50	62.52	218.92	250.16	156.21	56.27	-9.99	7.45	62.73	0.34	2.96	62.49	-0.04	3.23
	600	100	30.43	145.23	377.25	114.82	374.48	1130.62	344.98	30.70	0.88	2.92	30.43	-0.01	2.67
Γ_{12}	100	10	51.13	62.85	22.92	11.78	46.02	-9.99	6.15	53.72	5.06	4.30	52.30	2.28	3.82
	100	50	19.52	34.08	74.59	14.61	17.58	-9.95	2.36	20.18	3.36	2.38	20.97	7.43	1.46
	100	100	9.97	19.79	98.48	10.47	8.97	-10.01	2.36	10.21	2.44	2.81	11.07	11.01	4.17
	300	10	20.25	91.68	352.72	71.98	18.23	-9.99	5.27	20.48	1.15	0.41	26.80	32.32	7.06
	300	50	29.11	67.46	131.74	38.91	26.22	-9.94	3.08	29.19	0.28	0.12	29.46	1.19	3.77
	300	100	14.57	16.32	11.98	1.86	35.65	144.67	21.02	14.59	0.12	3.55	16.48	13.13	2.74
	600	10	28.53	24.26	-14.96	4.40	33.50	17.43	6.52	28.53	0.01	0.05	28.13	-1.39	3.23
	600	50	41.42	50.83	22.72	10.23	37.28	-10.00	4.59	41.53	0.25	1.48	42.43	2.44	1.72
	600	100	34.81	176.92	408.25	142.10	938.00	2594.63	903.14	35.02	0.61	1.61	35.71	2.59	2.12
Γ_2	100	10	34.01	41.27	21.34	7.37	30.61	-9.99	4.88	48.57	42.81	15.38	34.57	1.65	4.22
	100	50	45.08	61.82	37.13	16.75	40.59	-9.96	4.67	45.14	0.12	0.71	44.95	-0.29	0.73
	100	100	22.47	33.68	49.90	11.43	20.22	-10.03	2.36	22.98	2.25	4.17	23.69	5.45	2.30
	300	10	27.10	132.58	389.21	105.46	24.39	-9.99	2.72	27.60	1.85	1.24	36.34	34.10	9.79
	300	50	38.14	88.66	132.47	51.00	34.34	-9.96	3.89	38.38	0.63	0.61	38.39	0.65	4.80
	300	100	12.20	12.98	6.38	0.78	32.07	162.88	20.45	12.23	0.27	1.13	5.92	-51.48	6.30
	600	10	58.39	62.72	7.42	4.33	65.61	12.37	7.18	58.38	-0.02	2.80	58.38	-0.02	0.42
	600	50	28.24	116.47	312.43	88.52	25.41	-10.00	4.89	28.40	0.56	2.48	28.23	-0.04	0.68
	600	100	42.44	224.77	429.61	182.39	2412.57	5584.67	2370.27	42.61	0.41	3.11	41.74	-1.64	3.00

Table 8: Empirical mean, percentage of bias and square root of mean square error of the fixed effects' estimators regarding the first response variable according to correlations' variation.

Parameters	ρ	δ	θ_T	EM			Fisher scoring - EM			ECM			CMLME		
				$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE
$\beta_1^{(1)}$	0	0	15.27	14.99	-1.81	1.37	11.49	-24.78	4.59	11.48	-24.80	6.02	18.93	23.96	3.70
	5	0	5.31	5.22	-1.66	3.04	5.02	-5.53	1.07	5.20	-2.16	2.32	5.33	0.40	1.80
	95	0	11.38	11.25	-1.13	0.59	11.72	3.00	0.91	11.72	3.00	4.04	11.34	-0.38	3.16
	0	5	24.38	24.18	-0.82	1.24	23.85	-2.16	0.60	23.03	-5.53	2.62	24.41	0.14	3.28
	0	5	36.77	36.23	-1.47	2.48	27.03	-26.49	10.57	37.40	1.71	1.93	36.72	-0.15	3.82
	5	5	20.81	20.53	-1.37	0.87	16.08	-22.74	4.79	16.06	-22.84	6.45	20.79	-0.08	1.33
	50	5	31.33	31.07	-0.84	2.65	31.87	1.71	0.61	23.03	-26.49	9.00	31.34	0.04	1.08
	95	5	21.86	21.77	-0.40	1.27	22.04	0.82	3.81	27.72	26.82	5.88	21.85	-0.02	0.52
	0	50	27.57	27.28	-1.05	3.36	28.27	2.54	0.80	28.27	2.54	4.08	27.53	-0.14	0.59
	5	50	30.26	29.92	-1.12	0.36	26.44	-12.61	5.09	26.45	-12.61	4.08	30.22	-0.15	1.12
	50	50	13.51	13.32	-1.38	3.49	16.98	25.65	4.40	17.06	26.31	3.57	13.64	0.96	1.14
	95	50	16.10	15.79	-1.93	1.27	20.58	27.81	5.69	20.58	27.81	5.41	16.36	1.60	2.96
	0	95	35.68	35.57	-0.30	1.12	37.84	6.04	2.32	37.83	6.03	2.15	35.65	-0.09	0.97
	5	95	37.32	37.00	-0.85	1.54	40.70	9.06	4.26	40.70	9.06	4.84	37.25	-0.18	4.76
	50	95	16.63	16.36	-1.60	4.37	13.87	-16.58	5.28	13.87	-16.58	3.71	16.69	0.36	0.84
95	95	45.15	44.43	-1.61	3.94	54.37	20.42	10.65	54.37	20.42	9.55	45.33	0.40	3.02	
$\beta_2^{(1)}$	0	0	27.82	27.78	-0.14	3.00	20.03	-28.00	8.12	27.54	-1.00	1.14	27.71	-0.38	1.69
	5	0	38.85	38.83	-0.05	0.78	40.86	5.18	4.92	37.79	-2.74	4.58	38.77	-0.20	4.27
	50	0	18.26	18.27	0.07	3.39	17.77	-2.68	0.56	17.47	-4.33	4.81	18.23	-0.14	0.63
	95	0	10.25	10.25	0.01	3.30	10.15	-0.95	3.37	7.35	-28.29	4.76	10.26	0.08	1.37
	0	5	20.86	20.78	-0.37	2.89	19.28	-7.59	3.02	20.86	0.00	4.70	20.88	0.10	0.06
	5	5	49.43	49.30	-0.27	3.96	51.62	4.44	3.78	51.26	3.70	3.91	49.44	0.02	3.60
	50	5	11.06	11.04	-0.14	0.21	10.86	-1.80	4.41	10.25	-7.28	0.91	11.01	-0.45	1.40
	95	5	2.70	2.70	0.06	3.98	2.71	0.34	0.98	2.50	-7.59	4.66	2.68	-0.74	2.92
	0	50	30.46	30.46	0.00	4.76	28.39	-6.80	3.25	29.56	-2.95	4.31	30.48	0.07	4.40
	5	50	28.92	28.91	-0.05	3.15	27.94	-3.38	1.63	26.93	-6.89	3.19	28.91	-0.05	0.36
	50	50	19.69	19.68	-0.05	0.19	19.14	-2.79	3.28	17.04	-13.45	2.91	19.70	0.03	1.92
	95	50	38.63	38.60	-0.07	2.59	38.49	-0.36	1.86	18.25	-52.76	20.93	38.64	0.03	4.85
	0	95	5.23	5.22	-0.19	3.45	6.03	15.30	2.06	3.56	-31.93	4.66	5.22	-0.17	4.32
	5	95	2.30	2.29	-0.22	3.55	2.07	-10.13	0.47	2.64	14.83	3.89	2.29	-0.38	0.66
	50	95	17.83	17.78	-0.30	2.71	19.42	8.90	1.66	16.16	-9.38	4.72	17.90	0.38	0.63
95	95	5.28	5.30	0.38	2.91	5.57	5.50	0.78	5.75	8.90	0.49	5.30	0.47	3.92	
$\beta_3^{(1)}$	0	0	18.81	18.70	-0.58	0.25	21.38	13.66	4.27	21.94	16.64	3.15	18.75	-0.32	4.18
	5	0	42.00	41.90	-0.25	2.82	41.07	-2.22	4.22	41.90	-0.24	0.79	41.89	-0.27	0.80
	50	0	3.16	3.16	0.07	1.41	3.10	-1.79	4.59	3.13	-0.95	3.52	3.16	-0.06	3.02
	95	0	16.17	16.18	0.05	3.71	16.18	-0.17	4.83	15.80	-2.29	4.22	16.17	0.00	2.95
	0	5	14.89	14.81	-0.54	0.93	11.81	-20.67	4.07	19.60	31.63	4.83	14.81	-0.54	1.89
	5	5	17.61	17.57	-0.21	2.54	19.12	8.57	1.53	13.97	-20.67	4.70	17.65	0.23	3.55
	50	5	32.51	32.47	-0.11	3.97	33.82	4.03	4.48	33.95	4.43	2.59	32.56	0.15	3.21
	95	5	9.45	9.44	-0.09	1.45	9.49	0.42	2.77	9.44	-0.11	0.82	9.45	0.00	3.52

Table 8 - continued on next page

Table 8 - continued from previous page

Parameters	ρ	δ	θ_T	EM				Fisher scoring - EM				ECM				CMLME			
				$\hat{\theta}$		RMSE		$\hat{\theta}$		RMSE		$\hat{\theta}$		RMSE		$\hat{\theta}$		RMSE	
				Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE				
	0	50	29.91	29.93	0.08	0.13	24.97	-16.63	5.24	31.02	3.71	3.57	29.81	-0.33	3.96				
	5	50	35.65	35.60	-0.14	3.70	38.79	8.81	3.24	38.68	8.50	5.37	35.60	-0.14	0.94				
	50	50	7.11	7.10	-0.20	5.32	7.35	3.38	4.08	5.46	-23.16	4.25	7.10	-0.14	1.53				
	95	50	30.56	30.63	0.23	3.56	30.55	-0.03	5.23	23.34	-23.63	8.83	30.57	0.03	3.54				
	0	95	2.57	2.57	0.00	0.12	2.06	-19.76	0.76	1.76	-31.52	1.30	2.57	0.00	4.88				
	5	95	41.76	41.80	0.09	4.19	33.65	-19.42	9.38	33.53	-19.71	8.85	41.73	-0.07	3.83				
	50	95	38.77	38.81	0.10	0.63	34.55	-10.89	4.79	31.29	-19.30	7.53	38.72	-0.13	2.50				
	95	95	34.98	35.03	0.14	1.64	38.34	9.61	3.86	38.35	9.63	4.78	35.02	0.11	3.22				
$\beta_4^{(1)}$	0	0	46.79	46.67	-0.26	0.53	46.06	-1.55	4.47	46.06	-1.55	2.70	46.57	-0.48	5.35				
	5	0	10.35	10.33	-0.19	2.65	10.57	2.14	5.03	10.61	2.48	0.58	10.36	0.14	2.84				
	50	0	23.74	23.70	-0.15	2.16	22.65	-4.59	1.67	22.73	-4.25	3.18	23.77	0.11	3.68				
	95	0	36.65	36.61	-0.11	1.54	31.92	-12.91	5.77	31.87	-13.04	4.79	36.66	0.04	1.15				
	0	5	3.50	3.53	0.86	1.30	3.31	-5.56	1.65	3.17	-9.43	4.39	3.53	0.86	3.44				
	5	5	7.12	7.19	0.98	1.40	6.63	-6.86	4.32	7.73	8.61	4.95	7.15	0.42	1.00				
	50	5	28.44	28.48	0.14	3.05	39.22	37.92	10.87	30.80	8.31	4.02	28.50	0.21	1.73				
	95	5	7.55	7.56	0.13	1.02	13.86	83.54	6.75	7.06	-6.49	2.60	7.56	0.13	4.53				
	0	50	19.03	19.03	0.00	3.34	28.13	47.84	9.70	24.07	26.46	5.84	19.08	0.26	2.74				
	5	50	32.40	32.41	0.03	1.73	27.80	-14.20	4.80	38.19	17.87	7.95	32.38	-0.06	4.24				
	50	50	41.40	41.38	-0.05	0.05	43.99	6.25	5.30	35.53	-14.18	7.25	41.39	-0.02	3.39				
	95	50	18.28	18.21	-0.37	2.68	18.24	-0.22	1.28	16.00	-12.47	4.22	18.28	0.01	5.04				
	0	95	21.80	21.82	0.07	0.23	25.10	15.14	3.58	25.39	16.47	4.71	21.77	-0.15	2.57				
	5	95	40.93	40.99	0.15	3.34	34.02	-16.88	7.48	34.10	-16.68	6.89	40.89	-0.10	4.57				
	50	95	13.35	13.31	-0.28	3.03	16.11	20.67	2.82	16.14	20.90	2.79	13.34	-0.09	2.88				
	95	95	27.08	27.16	0.29	0.69	33.89	25.15	6.82	33.89	25.15	7.88	27.09	0.04	3.44				

Table 9: Empirical mean, percentage of bias and square root of mean square error of the fixed effects' estimators regarding the second response variable according to correlations' variation.

Parameters	ρ	δ	θ_T	EM			Fisher scoring - EM			ECM			CMLME		
				Bias	RMSSE	$\hat{\theta}$	Bias	RMSSE	$\hat{\theta}$	Bias	RMSSE	$\hat{\theta}$	Bias	RMSSE	$\hat{\theta}$
$\beta_1^{(2)}$	0	0	29.23	28.54	2.01	45.47	55.57	16.24	33.70	15.31	4.65	29.42	0.64	3.87	
	5	0	30.26	29.93	4.77	24.29	-19.72	5.98	31.12	2.83	3.50	30.17	-0.30	4.09	
	50	0	7.99	7.91	4.50	9.21	15.29	3.38	12.43	55.57	6.41	7.99	-0.05	4.86	
	95	0	29.06	28.92	1.56	29.88	2.83	0.85	23.33	-19.72	5.93	29.05	-0.04	2.38	
	0	5	5.42	5.03	0.41	0.75	-86.19	4.71	-4.67	-186.19	10.14	5.38	-0.75	0.50	
	5	5	11.30	10.59	0.74	3.57	-68.44	8.79	6.23	-44.90	6.14	11.24	-0.56	1.77	
	50	5	4.98	4.90	1.73	2.75	-44.71	3.32	1.57	-68.44	3.45	4.97	-0.11	5.16	
	95	5	47.60	47.44	1.83	47.15	-0.95	4.48	47.15	-0.95	1.00	47.59	-0.01	2.77	
	0	50	11.31	11.16	3.47	18.57	64.18	7.86	18.57	64.18	7.73	11.30	-0.08	0.26	
	5	50	10.42	10.27	3.00	4.89	-53.09	6.60	4.89	-53.09	6.40	10.40	-0.19	0.60	
	50	50	20.56	20.22	2.57	24.06	17.00	4.24	24.06	17.00	3.99	20.52	-0.21	3.66	
	95	50	24.28	23.81	2.45	25.37	4.50	3.48	25.40	4.61	3.36	24.17	-0.46	0.43	
	0	95	11.41	11.35	0.65	11.97	4.87	3.98	17.18	50.61	6.47	11.40	-0.05	0.64	
	5	95	30.99	30.10	3.05	19.98	-35.51	11.61	32.50	4.87	1.66	30.74	-0.80	0.93	
	50	95	3.29	3.13	4.91	4.96	50.69	1.90	2.12	-35.52	4.51	3.25	-1.11	2.29	
95	95	5.06	4.60	0.98	8.17	61.53	3.25	8.17	61.55	3.22	5.28	4.35	1.36		
$\beta_2^{(2)}$	0	0	35.51	36.00	1.37	55.51	56.31	21.03	33.44	-5.83	3.39	35.57	0.17	2.16	
	5	0	2.93	2.91	3.97	1.58	-46.14	1.67	4.57	55.97	4.48	2.93	0.05	1.01	
	50	0	5.70	5.72	3.69	4.91	-13.92	2.81	2.98	-47.72	2.73	5.70	0.00	3.29	
	95	0	19.40	19.41	0.06	4.69	-4.98	1.68	14.80	-23.71	4.77	19.40	0.00	3.46	
	0	5	46.26	47.10	1.82	0.92	-2.46	1.16	45.02	-2.68	4.26	46.24	-0.04	4.77	
	5	5	20.33	20.34	0.06	3.90	-4.43	3.71	17.25	-15.14	4.41	20.28	-0.25	1.20	
	50	5	22.13	22.14	0.05	2.54	-15.23	4.43	18.51	-16.36	5.61	22.24	0.50	2.29	
	95	5	1.65	1.65	1.61	1.34	-18.79	2.30	1.36	-17.71	4.12	1.64	-0.61	0.35	
	0	50	41.20	41.21	0.03	4.33	18.04	-56.21	23.31	18.07	-56.14	23.49	41.21	0.02	0.71
	5	50	47.64	47.61	-0.06	4.90	42.07	-11.69	7.14	41.94	-11.96	47.63	-0.03	1.63	
	50	50	32.33	32.25	-0.24	3.38	32.21	-0.37	3.45	32.12	-0.65	4.79	32.36	0.08	2.58
	95	50	14.00	14.06	0.43	2.59	13.97	-0.21	3.15	9.25	-33.93	13.98	-0.14	2.46	
	0	95	40.66	40.68	0.05	2.16	42.59	4.75	3.89	36.11	-11.19	40.67	0.02	1.03	
	5	95	30.93	30.91	-0.05	1.90	29.80	-3.64	2.90	32.38	4.69	30.95	0.07	3.63	
	50	95	28.53	28.55	0.07	4.67	-1.30	0.43	0.43	28.20	-1.16	3.76	28.56	0.10	4.83
95	95	37.28	37.25	-0.08	4.14	37.44	0.43	3.70	37.44	0.43	3.76	37.34	0.16	0.19	
$\beta_3^{(2)}$	0	0	38.57	38.61	0.10	0.09	39.70	2.93	34.98	-9.31	5.82	38.58	0.03	1.76	
	5	0	39.60	39.51	-0.23	0.45	38.24	-3.43	40.74	2.88	1.45	39.62	0.05	4.28	
	50	0	19.55	19.60	0.26	2.62	17.97	-8.06	3.86	18.83	-3.68	19.58	0.15	0.53	
	95	0	7.59	7.57	0.64	0.64	5.96	-21.48	5.88	-22.53	3.35	7.57	-0.26	3.06	
	0	5	29.32	29.12	-0.68	1.29	18.25	-37.76	11.37	45.26	16.53	28.98	-1.16	3.67	
	5	5	10.32	10.34	0.21	2.04	11.78	14.15	1.52	6.39	-38.06	10.33	0.14	4.23	
	50	5	43.84	43.92	0.19	0.38	41.09	-6.27	4.41	41.15	-6.14	43.87	0.07	3.85	
	95	5	21.08	21.05	-0.14	5.01	21.09	0.05	2.50	20.97	-0.52	21.07	-0.05	1.95	

Table 9 - continued on next page

Table 9 - continued from previous page

Parameters	ρ	δ	θ_T	EM			Fisher scoring - EM			ECM			CMLME		
				$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE
$\beta_4^{(2)}$	0	50	9.56	9.55	-0.09	2.98	21.78	127.82	13.39	21.90	129.08	12.42	9.43	-1.33	0.36
	5	50	6.78	6.77	-0.17	2.07	11.25	65.93	4.61	11.21	65.34	5.67	6.74	-0.52	4.39
	50	50	33.41	33.27	-0.42	0.70	32.25	-3.47	3.74	32.21	-3.59	2.43	33.45	0.12	3.97
	95	50	45.98	44.96	-2.21	1.31	45.94	-0.09	4.42	44.37	-3.50	4.00	46.04	0.13	2.65
	0	95	42.14	42.14	0.00	3.84	41.34	-1.90	1.77	40.12	-4.79	4.46	42.14	0.00	3.04
	5	95	7.77	7.76	-0.07	4.13	7.39	-4.86	4.20	7.32	-5.76	3.50	7.77	0.02	4.55
	50	95	45.11	45.08	-0.07	1.89	42.13	-6.61	3.39	42.25	-6.34	3.50	45.12	0.03	3.75
	95	95	13.88	13.84	-0.29	3.48	14.93	7.56	2.63	14.97	7.85	2.33	13.87	-0.07	2.81
	0	0	4.84	4.81	-0.62	0.65	5.04	4.13	2.00	3.47	-28.31	2.88	4.76	-1.65	3.45
	5	0	43.30	43.26	-0.09	5.06	32.83	-24.17	10.70	45.11	4.18	2.40	43.22	-0.18	0.83
	50	0	36.82	36.79	-0.08	3.02	18.76	-49.05	18.34	18.71	-49.19	18.10	36.87	0.14	4.81
	95	0	9.63	9.64	0.10	3.13	-14.29	-248.39	23.93	-14.23	-247.77	23.88	9.64	0.12	0.99
	0	5	31.90	31.36	-1.69	2.80	38.29	20.04	6.40	39.12	22.63	7.52	31.89	-0.03	2.00
	5	5	36.10	36.17	0.19	0.79	41.64	15.36	5.68	43.31	19.96	7.70	36.18	0.23	4.61
	50	5	48.65	48.62	-0.06	0.89	55.94	14.98	8.33	56.38	15.89	8.59	48.49	-0.33	2.44
95	5	7.68	7.68	0.04	0.11	8.25	7.37	1.86	8.80	14.55	1.51	7.75	0.91	0.32	
0	50	4.43	4.46	0.68	3.87	18.37	314.67	14.52	18.32	313.54	13.86	4.33	-2.26	2.30	
5	50	9.78	9.74	-0.38	1.35	32.07	227.91	23.04	32.27	229.96	22.45	9.73	-0.54	3.29	
50	50	18.44	18.40	-0.23	4.28	16.48	-10.63	2.01	16.46	-10.74	1.98	18.40	-0.20	3.75	
95	50	21.79	21.81	0.10	1.60	21.78	-0.05	2.25	21.00	-3.63	3.68	21.77	-0.09	2.80	
0	95	2.96	2.96	-0.03	1.79	11.70	285.27	9.26	24.11	714.53	21.57	2.93	-1.01	4.80	
5	95	4.94	4.93	-0.20	1.19	-4.68	-194.74	9.74	-4.77	-196.56	9.78	4.92	-0.42	2.27	
50	95	7.22	7.19	-0.42	4.70	8.30	14.96	3.56	8.24	14.13	1.15	7.21	-0.20	2.37	
95	95	34.86	34.71	-0.43	0.87	36.97	6.05	5.35	36.96	6.02	2.17	34.87	0.03	4.05	

Table 10: Empirical mean, percentage of bias and square root of mean square error of the estimators of residuals' matrix of covariance according to correlations' variation.

Parameters	ρ	δ	θ_T	EM			Fisher scoring - EM			ECM			CMLME			
				$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE	
σ_1^2	0	0	4.28	4.84	13.07	1.84	8.67	102.56	4.75	6.26	46.15	4.56	4.28	-0.06	0.48	
	5	0	1.35	1.37	1.63	4.47	1.38	2.56	1.64	1.57	16.20	3.74	1.35	-0.10	0.75	
	50	0	4.67	4.72	1.10	2.74	5.68	21.62	2.16	5.56	19.04	0.89	4.64	-0.56	1.88	
	95	0	3.10	3.11	0.29	1.46	7.06	127.73	4.04	5.04	62.52	2.01	3.30	6.43	3.94	
	0	5	8.58	8.62	0.42	3.48	525.33	6022.70	517.14	12.62	47.12	4.03	8.50	-0.95	0.39	
	5	5	24.21	24.29	0.33	3.57	179.10	639.77	155.33	33.18	37.06	9.29	24.77	2.29	4.52	
	50	5	19.18	19.24	0.29	1.52	31.79	65.76	12.96	29.29	52.71	10.15	19.18	-0.01	3.25	
	95	5	8.01	8.01	0.04	4.98	8.56	6.92	0.56	13.72	71.28	5.67	7.99	-0.24	4.97	
	0	50	6.20	6.30	1.54	1.84	46.36	647.73	40.61	8.45	36.32	2.71	6.68	7.67	2.48	
	5	50	20.88	21.03	0.74	4.77	23.35	11.81	2.60	34.70	66.17	14.15	21.55	3.23	3.28	
	50	50	9.30	9.35	0.58	2.69	9.37	0.78	3.71	18.08	94.38	8.87	9.23	-0.71	3.33	
	95	50	15.05	15.09	0.29	3.43	15.05	0.00	2.46	39.83	164.63	24.77	13.78	-8.44	1.90	
	0	95	3.17	3.17	0.14	2.66	3.29	3.65	1.51	3.85	21.38	5.06	3.17	-0.01	1.56	
	5	95	20.34	20.44	0.48	4.26	23.04	13.25	5.08	22.08	8.56	2.68	21.29	4.65	4.16	
	50	95	6.55	6.59	0.68	2.42	17.24	163.26	11.48	7.65	16.76	2.33	5.29	-19.25	1.65	
	95	95	19.27	19.67	2.07	2.89	56.13	191.30	37.02	30.74	59.54	11.48	19.26	-0.04	4.32	
	σ_{12}^2	0	0	0.00	0.00	-	4.33	0.00	-	0.57	0.00	-	2.36	-	-	-
		5	0	0.00	0.00	-	3.95	0.00	-	2.57	0.00	-	2.88	-	-	-
50		0	0.00	0.00	-	1.35	0.00	-	5.06	0.00	-	4.78	-	-	-	
95		0	0.00	0.00	-	1.06	0.00	-	2.63	0.00	-	2.69	-	-	-	
0		5	0.68	0.74	8.96	1.55	-96.00	-14218.08	96.76	0.56	-17.39	2.70	-	-	-	
5		5	0.43	0.47	9.63	3.05	60.22	13904.28	59.77	0.64	49.71	0.65	-	-	-	
50		5	0.85	0.94	10.27	1.93	51.42	5949.53	50.58	1.56	83.30	3.34	-	-	-	
95		5	0.16	0.18	14.79	3.00	-1.21	-855.63	2.27	0.02	-85.53	0.63	-	-	-	
0		50	5.47	5.50	0.57	2.92	426.29	7693.32	421.38	6.19	13.11	1.75	-	-	-	
5		50	2.54	2.57	1.03	3.13	3.73	46.66	3.65	3.76	47.86	1.41	-	-	-	
50		50	7.20	7.37	2.43	2.08	7.24	0.61	2.69	12.33	71.22	7.08	-	-	-	
95		50	3.08	3.21	4.08	0.79	3.08	-0.02	3.78	5.71	85.34	2.67	-	-	-	
0		95	6.66	6.74	1.24	1.82	6.97	4.63	1.73	8.84	32.73	5.40	-	-	-	
5		95	15.96	16.07	0.69	2.15	17.10	7.15	3.05	23.18	45.25	7.79	-	-	-	
50		95	11.77	11.84	0.64	2.71	24.81	110.77	13.08	17.37	47.61	5.97	-	-	-	
95		95	8.88	8.93	0.54	3.11	24.61	177.11	15.75	15.07	69.76	6.19	-	-	-	
σ_2^2		0	0	4.62	4.77	3.24	1.89	10.62	129.79	7.41	7.30	58.06	2.86	5.44	17.79	1.63
		5	0	5.52	5.65	2.44	1.48	11.40	106.57	6.81	6.65	20.54	3.11	4.67	-15.46	2.24
	50	0	9.49	9.58	0.97	1.74	324.13	3315.45	314.45	12.51	31.85	3.12	9.74	2.67	3.60	
	95	0	1.59	1.60	0.40	3.41	99.87	6181.04	98.47	2.16	35.64	3.97	1.59	-0.26	1.03	
	0	5	21.72	21.84	0.58	1.76	1690.97	7685.32	1669.38	102.59	372.33	80.90	21.34	-1.74	3.59	
	5	5	3.10	3.13	1.00	4.30	151.03	4771.90	147.23	11.74	278.64	8.62	3.22	3.98	2.38	
	50	5	14.98	15.15	1.14	3.01	54.61	264.53	39.79	51.56	244.19	36.74	15.93	6.33	3.28	
	95	5	1.30	1.33	2.57	3.44	3.10	138.32	1.84	4.07	212.70	4.44	1.52	16.74	0.32	
	0	50	19.27	19.32	0.24	4.00	74.18	284.95	55.03	54.10	180.72	35.30	19.27	-0.02	1.17	
	5	50	1.23	1.24	0.49	4.64	1.29	4.88	0.91	3.06	148.58	2.78	1.22	-0.43	4.08	
	50	50	22.28	23.56	5.75	3.92	22.32	0.18	4.47	39.07	75.37	16.96	23.02	3.31	2.04	
	95	50	2.53	2.70	6.75	1.06	2.53	0.12	4.29	3.86	52.53	1.46	2.53	-0.03	2.25	
	0	95	15.52	15.55	0.20	3.73	16.06	3.48	5.10	65.47	321.82	50.03	16.74	7.85	2.12	
	5	95	13.25	13.35	0.73	4.41	13.95	5.29	1.88	44.68	237.23	31.87	13.19	-0.46	3.67	
	50	95	23.43	23.70	1.14	2.25	39.32	67.83	15.62	40.07	71.02	16.78	23.42	-0.05	3.55	
	95	95	4.54	4.60	1.43	0.97	12.75	180.83	8.29	5.67	24.93	2.35	3.59	-20.82	0.95	

Table 11: Empirical mean, percentage of bias and square root of mean square error of the estimators of random effects' matrix of covariance according to correlations' variation.

Parameters	ρ	δ	θ_T	EM			Fisher scoring - EM			ECM			CMLME			
				$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE	$\hat{\theta}$	Bias	RMSE	
Γ_1	0	0	46.93	272.52	480.70	225.67	44.87	-4.40	2.92	47.08	0.33	2.75	46.92	-0.01	2.60	
	5	0	4.03	6.48	60.76	5.73	3.63	-9.98	1.79	4.47	10.81	0.44	4.02	-0.25	2.79	
	50	0	21.33	22.28	4.47	2.12	19.20	-9.99	3.50	21.35	0.09	2.69	21.26	-0.31	3.08	
	95	0	26.45	26.88	1.61	0.91	23.80	-10.03	3.22	26.52	0.25	0.12	26.55	0.39	4.04	
	0	5	59.03	76.04	28.81	17.30	53.12	-10.02	6.45	59.04	0.02	1.69	58.31	-1.23	3.63	
	5	5	27.01	30.33	12.29	5.82	37.56	39.06	10.55	27.26	0.93	3.93	27.12	0.42	3.02	
	50	5	16.87	18.23	8.04	4.09	15.18	-10.01	3.12	16.92	0.31	1.74	16.82	-0.30	1.14	
	95	5	19.09	19.76	3.53	3.98	17.19	-9.97	1.91	19.09	0.00	1.26	19.09	-0.01	2.38	
	0	50	45.86	78.39	70.92	32.79	41.27	-10.00	4.90	45.94	0.18	2.60	48.22	5.14	3.36	
	5	50	11.47	21.97	91.58	11.15	12.62	10.02	4.77	11.53	0.48	2.68	12.30	7.25	4.71	
	50	50	30.81	134.24	335.69	103.96	27.72	-10.03	3.09	31.27	1.50	0.69	27.30	-11.40	4.01	
	95	50	18.54	87.95	374.40	69.34	23.27	25.52	4.76	19.11	3.09	4.47	8.76	-52.75	10.05	
	0	95	8.76	9.73	11.05	4.81	8.39	-4.20	0.51	8.77	0.09	3.19	9.74	11.15	4.38	
	5	95	22.76	49.64	118.12	26.63	20.48	-10.03	4.58	23.14	1.66	3.86	22.34	-1.84	0.43	
	50	95	20.90	46.46	122.32	25.77	18.81	-10.01	2.43	21.28	1.82	3.08	20.82	-0.38	1.81	
	95	95	62.52	218.92	250.16	156.39	56.27	-9.99	7.64	62.73	0.34	2.42	62.49	-0.04	4.46	
	Γ_{12}	0	0	0.00	0.00	-	0.31	0.00	-	4.75	0.00	-	3.01	0.00	-	1.18
		5	0	0.50	22.18	4335.58	21.63	0.45	-9.40	0.73	0.84	67.21	2.24	0.64	28.01	4.41
		50	0	12.85	32.31	151.48	19.47	11.57	-9.99	1.68	12.88	0.25	4.47	14.76	14.85	4.64
		95	0	31.24	34.82	11.47	5.24	28.11	-10.02	3.35	31.26	0.05	1.86	31.59	1.13	3.71
0		5	0.00	0.00	-	2.29	0.00	-	0.53	0.00	-	4.27	0.00	-	3.89	
5		5	3.11	9.33	200.02	7.77	4.18	34.47	1.70	3.24	4.13	2.30	3.72	19.63	3.28	
50		5	5.98	1.27	-78.68	4.73	5.38	-9.97	1.78	6.19	3.44	2.52	5.53	-7.57	4.77	
95		5	31.74	35.27	11.11	5.39	28.57	-9.99	3.17	31.75	0.03	0.05	32.09	1.09	0.68	
0		50	0.00	0.00	-	1.62	0.00	-	0.69	0.00	-	3.93	0.00	-	2.12	
5		50	1.45	7.63	426.32	7.18	1.32	-9.10	2.40	1.63	12.54	3.51	1.34	-7.32	1.50	
50		50	21.14	94.95	349.15	73.84	19.02	-10.05	5.20	21.50	1.68	1.89	29.98	41.80	9.71	
95		50	17.93	3.96	-77.90	14.50	27.54	53.58	10.22	18.02	0.48	3.99	24.07	34.27	6.13	
0		95	0.00	0.00	-	1.01	0.00	-	4.94	0.00	-	0.21	0.00	-	2.03	
5		95	0.90	23.45	2505.61	23.37	0.81	-10.33	0.99	0.98	8.78	2.95	1.31	45.72	0.46	
50		95	14.54	73.90	408.25	59.17	13.08	-10.01	1.45	15.00	3.16	2.36	20.36	40.02	6.00	
95		95	41.42	50.83	22.72	9.38	37.28	-10.00	4.12	41.53	0.25	2.86	42.43	2.44	1.27	
Γ_2		0	0	29.31	128.88	339.72	99.57	29.98	2.29	3.63	29.77	1.56	4.59	37.40	27.61	8.56
		5	0	21.80	56.36	158.53	34.41	19.62	-10.02	4.64	22.08	1.30	3.63	19.54	-10.36	5.19
		50	0	18.46	39.26	112.67	20.81	16.71	-9.45	2.43	18.56	0.52	3.35	21.03	13.90	3.67
		95	0	38.59	49.65	28.67	11.90	34.73	-10.00	6.22	38.59	0.01	0.90	38.55	-0.12	3.11
	0	5	58.61	155.92	166.03	97.33	52.75	-10.00	6.81	59.39	1.33	3.01	63.48	8.32	4.94	
	5	5	66.75	125.11	87.43	57.82	78.05	16.92	11.32	66.79	0.06	2.21	65.67	-1.62	4.59	
	50	5	7.33	12.22	66.70	4.89	6.60	-9.99	3.42	7.37	0.48	4.77	8.32	13.53	1.01	
	95	5	56.19	64.27	14.39	8.31	50.57	-10.00	6.52	56.21	0.03	2.41	56.95	1.35	4.70	
	0	50	35.71	35.90	0.52	3.15	32.15	-9.98	3.54	35.82	0.32	4.46	35.69	-0.05	1.33	
	5	50	37.98	74.10	95.12	36.12	34.18	-10.02	3.86	38.15	0.44	2.18	37.98	-0.01	3.52	
	50	50	18.53	82.29	344.10	64.47	16.69	-9.93	2.35	18.84	1.66	5.04	24.99	34.88	7.56	
	95	50	18.32	87.99	380.32	69.82	19.13	4.40	1.45	18.73	2.26	4.03	18.32	-0.02	0.92	
	0	95	37.49	44.12	17.68	6.67	37.25	-0.63	2.75	38.25	2.03	3.14	3.58	-90.45	34.49	
	5	95	8.33	27.55	230.71	19.21	7.50	-10.01	1.44	8.43	1.21	3.07	10.73	28.80	3.65	
	50	95	26.65	104.58	292.43	77.89	23.98	-10.03	5.38	26.91	0.99	0.64	24.99	-6.25	3.06	
	95	95	28.24	116.47	312.43	88.26	25.41	-10.00	4.52	28.40	0.56	1.95	28.23	-0.04	4.82	