AJAS / SPAS

African Journal of Applied Statistics
Vol. 5 (2), 2018, pages 489 – 502.
DOI: http://dx.doi.org/10.16929/ajas/489.226

ISSN 2316-0861

Fitting an optimal variance-covariance structure for longitudinal data under Linear Mixed Effects Models framework: simulation based analysis

Aubin Amagnide[1,2,*], **Micheline Gbeha**[3] **and Romain Glèlè Kakaï**[1]

[1]Laboratoire de Biomathématiques et d'Estimations Forestières (LABEF), Faculté des Sciences Agronomiques, Université d'Abomey-Calavi, 04 BP 1525, Cotonou, République du Bénin. [2] Programme Information Scientifique et Biométrie (PIS-B), Centre de Recherches Agricoles d'Agonkanmey (CRA-Agonkanmey), Institut National des Recherches Agricoles du Bénin (INRAB), 01 BP 884, Cotonou, République du Bénin.
[3]Département de Mathématiques, Faculté des Sciences et Techniques, Université d'Abomey-Calavi, République du Bénin.

Received : January 12, 2018. Accepted : August 01, 2018. Published Online : October 17, 2018.

Abstract. This paper assessed the performance of 5 fit statistics to determine the correct within subject covariance structure (WSCS) in longitudinal data analysis and investigated the consequence of misspecification of WSCS using Monte Carlo procedure ... (See page 490 for the full Abstract).

Key words: Repeated measurements; within-subject covariance structure; fit statistics; misspecification; Monte Carlo experiments.
AMS 2010 Mathematics Subject Classification : 97K80; 78M31; 68U20.

---

*Corresponding author Aubin Amagnide: aubin.amagnide@gmail.com
Micheline Gbeha : micgbeha@yahoo.fr
Romain Glèlè Kakaï : glele.romain@gmail.com

Full Abstract. In this study, we (i) assessed the performance of 5 fit statistics (AIC, BIC, HQIC, CAIC and AICC) to determine the cor-rect within-subject covariance structure (WSCS) in longitudinal data analysis and (ii) investigated the consequence of misspecification of WSCS. Firstly, a simulation study was achieved in 192 cases taking into account six characteristics of the data sample (sample size, measurement periods, magnitude of growth parameter, size of G matrices, covariance structure and distribution of the within-subject error). For each combination of these parameters, 500 replications were generated using Monte Carlo procedure and the hit rate of each of the 5 search statistics is computed and help to compare their performance. At a second step, based on 32 restricted simulation conditions, the effect of misspecification in WSCS was assessed by computing the mean relative bias and mean relative errors of the coefficients of fixed effects and random components. Results showed an overall best performance of the HQIC, BIC and CAIC for searching first order auto-regressive [AR(1)] and first order moving average [MA(1)] covariance structures.

Résumé. Dans la présente étude, (i) la performance de 5 critères d'information statistique (AIC, BIC, HQIC, CAIC et AICC) dans la détermination de la matrice de covariance entre mesures répétées dans l'analyse des données longitudinales et (ii) les conséquences d'une mauvaise spécification de la matrice de covariance entre mesures répétées ont été évaluées. Premièrement, une simulation a été réalisée dans 192 situations déterminées par les caractéristiques de l'échantillon de données (taille de l'échantillon, nombre de mesures répétées, paramètre de croissance, taille de la matrice G, matrice de covariance entre mesures répétées et la distribution des erreurs entre mesures répétées). Pour chaque combinaison de ces paramètres, le taux de succès de chaque critère d'information statistique est calculé dans le but de comparer les performances des 5 critères d'information statistique. Deuxièmement, sur la base de 32 situations restreintes déterminées par les caractéristiques de l'échantillon de données, l'effet d'une mauvaise spécification de la matrice de covariance entre mesures répétées a été évalué par la détermination des écarts et biais relatifs moyens des effets fixes et aléatoires estimés. Les résultats obtenus de la simulation montrent de meilleures performances globales pour HQIC, BIC and CAIC dans l'identification des matrices de covariance autoregressive de premier ordre [AR(1)] et de moyenne mobile de premier ordre [MA(1)]. Concernant la matrice de covariance autoregressive à moyenne mobile de premier ordre [ARMA(1,1)], les critères AIC, AICC et HQIC présentent les meilleures performances globales. Les résultats obtenus montrent également que, quelle que soit la situation de simulation considérée, les effets fixes étaient bien estimés avec cependant, une tendance au biais lorsque le paramètre de croissance tend à devenir petit. Par contre, les effets aléatoires étaient mal estimés au regard du biais relatif. Pour une bonne estimation des effets aléatoires, une attention particulière doit être accordée à la recherche de la matrice de covariance entre mesures répétées optimale dans l'analyse des données longitudinales.

1. Introduction

Longitudinal data (LD) constitutes a hierarchical struc-ture, with repeated observations over time nested within individuals (Steele (2008)). Because standard statistical models fail to recognize hierarchical structure, they become inappropriate methods to deal with these types of data (Snijders and Bosker (1999), Maas and Hox (2004)). Con-trary to standard statistical models, linear mixed effects models (LMEM) recognize the existence of such data hierarchies by allowing for residual components at each level in the hierarchy. Therefore, LMEM have widely been used to analyze LD where the measurement occasions are nested within cases (e.g. individual or subject) (Brandon (2013), AL-Marshadi (2014)).

In LD, observations are made at multiple time points on each subject. Thus, measures on the same subject at dif-ferent times tend to be correlated (McCulloch (2006)). Moreover, measures taken close together in time are more highly correlated than measures taken far apart in time (Hedeker and Gibbons (2006), Gibbons *et al.* (1979)). Hence, taking this dependency into account by specifying right covariance structure for observations within each subject becomes an important issue (Brandon (2013), AL-Marshadi (2014)). Specifically in longitudinal data analysis (LDA), information about change in the response variable over time is reflected only in the covariance matrix of the with-in-subject residuals (Hedeker and Gibbons (2006)). Some fit statistics are used to assess the adequacy of the covariance matrix structure considered according to the observed data (Yanosky II (2007), AL-Marshadi (2014)). These are Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Hannan and Quin Infor-mation Criterion (HQIC), Consistent Akaike Information Criterion (CAIC) and Akaike Information Criterion - Cor-rected (AICC) [Yanosky II (2007)]. Choosing an accurate criterion among those cited above constitutes an important issue for users of LMEM in LDA due to misleading results from covariance matrix misspeci-fication in statistical modeling (AL-Marshadi (2014)).

Moreover, although the LMEM allows for flexible mod-eling of LD, the simulation research literature is not nearly as extensive as standard methods. Few research works (Brandon (2013)) to date started exploring effect of mis-specification of within-subject error covariance. Unfortu-nately, apart from Brandon (2013), these studies were im-plemented under perfect model conditions (i.e. normally distributed random effects and residuals). However, it is known that real world data are rarely normally distributed and can deviate quite substantially from a Normal distribu-tion (Micceri (1989)). Therefore, this study aims to (i) assess the performance of 5 fit statistics in identifying the correct within-subject covariance structure in LDA and (ii) investi-gate the consequence of misspecification of within-subject covariance structure in LDA.

2. Methods

2.1. Model specification

Let us consider a simple linear growth model (here, a two-level growth model), written in matrix form as:

$$y = X\beta + Zu + \epsilon \tag{1}$$

where u and ϵ are assumed to be independently and identically distributed (multivariate normally distributed) with

$$\mathbb{E}\begin{bmatrix} u \\ \epsilon \end{bmatrix} = 0 \;\; and \;\; Var\begin{bmatrix} u \\ \epsilon \end{bmatrix} = \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix} \tag{2}$$

In equation (1), y is a column vector of repeated measures outcome, X is the known design matrix of fixed effects, β is a column vector of unknown fixed effect parameter estimates, Z is the known design matrix of the random effect, u is a column vector of unknown random effect parameter estimates and ϵ is the column vector of error associated with the measurement outcome. With longitudinal data, R and G correspond to the within-subject and between-subject error structures, respectively. Equation (1) can be rewritten in multilevel form as:

$$Level1 : y_{ti} = \beta_{0i} + \beta_{1i}Time + \epsilon_{ti}; \epsilon_{ti} \sim N(0, R) \tag{3}$$

$$Level2 : \beta_{0i} = \gamma_{00} + u_{0i}; \beta_{1i} = \gamma_{10} + u_{1i} \tag{4}$$

with

$$\begin{bmatrix} u_{0i} \\ u_{1i} \end{bmatrix} \sim N\left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, G = \begin{bmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{bmatrix}\right] \tag{5}$$

In (3), y_{ti} represents the individual trajectory of change; it is a function of time in level one with i standing for individual and t indicates time points index. The level two outcome variables β_{0i} and β_{1i} are the growth parameters in the model (see equation 4). β_{0i} and β_{1i} are multivariate normally distributed and vary around their grand means (γ_{00} and γ_{10}, respectively) with variances τ_{00} and τ_{11} respectively and covariance τ_{01}. The G matrix considered in (5) is called unstructured matrix.

2.2. Covariance structures and sit statistics considered

Four structures of the R matrix are considered in this study: the Independence structure, ID, the first order autoregres-sive structure, AR(1), the first order moving average struc-ture, MA(1) and the first order autoregressive moving av-erage model, ARMA(1,1). The simplest covariance matrix for R is the diagonal structure, known as independence structure. For 4 time points, we have:

$$R = \sigma^2 I = \begin{bmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{bmatrix} \tag{6}$$

where σ^2 is the constant variance (within subject error) at each time point.
The covariance matrix for AR(1) with four time points can be expressed as follow:

$$R = \sigma^2 \begin{bmatrix} 1 & \varphi_1 & \varphi_1^2 & \varphi_1^3 \\ \varphi_1 & 1 & \varphi_1 & \varphi_1^2 \\ \varphi_1^2 & \varphi_1 & 1 & \varphi_1 \\ \varphi_1^3 & \varphi_1^2 & \varphi_1 & 1 \end{bmatrix} \tag{7}$$

where φ_1 represents the correlation between two time points separated by a lag of one and σ^2, defined as above.
The form of covariance matrix for MA(1) with four time points can be expressed as follow:

$$R = \sigma^2 \begin{bmatrix} 1 & \theta_1 & \theta_1 & 0 \\ \theta_1 & 1 & \theta_1 & \theta_1 \\ \theta_1 & \theta_1 & 1 & \theta_1 \\ 0 & \theta_1 & \theta_1 & 1 \end{bmatrix} \tag{8}$$

where θ_1 represents the correlation between two time points separated by a lag of one and σ^2, defined as above.
The covariance matrix for ARMA(1,1) is as follow:

$$R = \sigma^2 \begin{bmatrix} 1 & \varphi_1\theta_1^0 & \varphi_1\theta_1^1 & \varphi_1\theta_1^2 \\ \varphi_1\theta_1^0 & 1 & \varphi_1\theta_1^0 & \varphi_1\theta_1^1 \\ \varphi_1\theta_1^1 & \varphi_1\theta_1^0 & 1 & \varphi_1\theta_1^0 \\ \varphi_1\theta_1^2 & \varphi_1\theta_1^1 & \varphi_1\theta_1^0 & 1 \end{bmatrix} \tag{9}$$

where σ^2, φ_1 and θ_1 are defined as above. If $\theta_1 = \varphi_1$, the structure becomes the one of ARMA(1,0) or just simply AR(1) defined above.

The fit statistics compared in the study are summarized in Table 1.

Table 1. Fits Statistics Considered and their Formula (L= likelihood of the model adjusted, p=number of parameters esti-mates, n=sample size and ln= natural logarithm.)

Fit statistics	Formula	Source
AIC	2p–2ln(L)	Akaike (1974)
BIC	pln(n)–2ln(L)	Schwarz (1978)
CAIC	p(ln(n)+1)–2ln(L)	Bozdogan (1987)
HQIC	2p(ln(ln(n)))–2ln(L)	Hannan nd Quinn (1979)
AICC	AIC+2np/(n–p–1)	Hurvich and Tsai (1989)

2.3. Simulation design and performance

2.3.1. Assessment of performance of the 5 fit statistics

Factors considered for the simulation are the sample size (50, 100, 150 and 200), the measurement periods (5 and 8), the magnitude of growth parameter (i.e. the mean of the individual slopes) so that $\beta_1 = 0.05$ and $\beta_1 = 0.16$, the size of G matrix

(small [$\tau_{00} = 0.1$ and $\tau_{11} = 0.05$] and medium [$\tau_{00} = 0.2$ and $\tau_{11} = 0.1$]) and the covariance structure (true R matrix) for generating the data [AR(1), MA(1) and ARMA(1,1)]. The sixth factor taken into account is the distribution of the within-subject error: Normal or Chi-square with 1 degree of freedom. Thus, a total of 192 combinations of factors have been considered. To avoid finding a single extreme data condition, five hundred replications were generated for each combination of factors using Monte Carlo procedure. Each dataset was then analyzed using four separate specifications of the R matrix (ID, AR(1), MA(1) and AR-MA(1,1)). Coefficient β_0 i.e. β_{00} was fixed to 0.10 for all combinations of factors. Three parameters were necessary to specify the three chosen error covariance structures: σ^2 (variance of the within subject errors), θ_1 (i.e., moving aver-age coefficient) and φ_1 (i.e., autoregressive correlation coefficient). The variance σ^2 was set as 2 and coefficients θ_1 and φ_1 were fixed to 0.50 and 0.8 (respectively).
The hit rate of each search statistics was used as the major criterion. A correct hit in model selection was represented by an event that the smallest fit index value for the hypothesized covariance structure matches the true covariance structure. Fit index hit rate for all investigated conditions and within-subject covariance structures was computed respectively. Moreover, the convergence rate of the analyses when specifying different R matrices regardless of the true R matrix was also computed. It is defined by an event that a model with a given R matrix specification converges.

2.3.2. Investigation of the consequence of misspecification in within-subject variance-covariance structure

The simulation used a total of 4 sample sizes (50, 100, 150 and 200) $\times$ 2 measurement occasions (5 and 8) $\times$ 2 (magnitude of growth parameter β_1: 0.05 or 0.16) $\times$ 2 size of G matrix (small [τ_{00}=0.1 and τ_{11}=0.05] and medium [τ_{00}=0.2 and τ_{11}=0.1]) with R matrix specified as AR(1) structure. This gave a total of 32 simulated data conditions. Five hundred replications were generated for each simulat-ed data condition, resulting in 32$\times$500 = 16,000 total da-tasets. These datasets were then analyzed for examining effect of under, over and generally misspecification in within subject variance-covariance structure (R). The other simulated data conditions (i.e. β_0, σ^2 and φ_1) were specified as previously.
Two criteria were used to examine the effects of mis-specification of the within subject covariance structure: (i) relative bias (RB) and relative error (RE) of the estimates of the fixed effects (i.e., intercept β_0 and slope β_1) and the random components (i.e., variances: τ_{00}, τ_{11} and $\sigma^2\epsilon$ and covariance: τ_{01}). Parameters RB and RE were calculated as follow:

$$RB = \frac{\hat{\gamma} - \gamma}{\gamma} \ \ and \ \ RE = \frac{|\ \hat{\gamma} - \gamma\ |}{\gamma} \tag{10}$$

In (10), γ is the true parameter value (i.e., β_0, β_1, τ_{00}, τ_{11}, $\sigma^2\epsilon$ and τ_{01}) and $\hat{\gamma}$ is the corresponding sample estimate. The mean relative bias (MRB) and the mean relative error (MRE) related to each estimator were computed for each of the 32 combinations of the factors.

The effect of design factors on MRB and MRE of the model parameters under different specifications of R matrix regardless of the true R matrix was assessed using an ANOVA and and statistical significance was set up at 0.1 % due to the large sample size and statistical power.

The MRE of the fixed effects and random components from the fitted models related to each combination of the factors are replaced by ranks. For a given combination of the factors, the ranks of the MRE are determined, the lowest MRE having the rank 1. The median ranks of the MRE are determined for some factor levels as well as for some groups of the factor levels based on the sample size and the measurement periods. The median rank of each of the 4 covariance structures) for all the 32 combinations of the factors is also computed.

3. Results

3.1. Performance of information statistics on searching for the correct within-subject covariance structure

3.1.1. Convergence rate

The average convergence rate (CR) of the estimation algorithm by the gen-erated and fitted covariance structures tended to be low ranging from a 25 % to 100 % (Table 2). Low CR tended to occur when the covariance structure was overspecified (e.g. ARMA(1,1) structure fit to an AR(1) structure) or when a generally misspecified covariance structure was fitted (e.g. MA(1) structure fit to a ARMA(1,1) structure or AR(1) structure fit to MA(1) structure). In general, the AR(1) and ARMA(1,1) fitted structures had the worst CR compared to the other fitted structures and the independent (ID) struc-ture had the best convergence rate, which was not surprising as no additional terms need to be estimated with an inde-pendent structure contrasting with MA(1) structure.

Table 2. Convergence rates (%) by generated and fitted covariance structure: mean and standard deviation (SD)

Generated covariance structure	Statistics	Fitted covariance structure			
		ID	AR(1)	MA(1)	ARMA(1,1)
AR(1)	Mean	100.00	93.68	99.86	38.39
	SD	0.00	9.23	0.28	7.70
MA(1)	Mean	85.06	24.96	33.79	31.52
	SD	20.15	11.89	16.55	14.52
ARMA(1,1)	Mean	100.00	90.56	99.98	35.99
	SD	0.00	13.39	0.08	6.94

3.1.2. Information Statistics Performance

For AR(1) covariance structure, the results of ANOVA conducted on fit statistics hit rates to investigate the impact of design factors reveal that all fit statistics hit rates were significantly affected by measurement periods and G matrix, except HQIC for which, only G matrix has significant effects (not presented). Moreover, the interaction between both factors were significant, meaning that the observed difference between measurement periods depend on G ma-trix and vice-versa. From the mean values of fit statistics performance (Table 3), the lowest values of BIC, CAIC and HQIC hit rate were found for 5 measurement periods while the lowest values of AIC and AICC hit rate were found for 8 measurement periods. Regarding the G matrix, the highest values for all hit rates were found for small size of G matrix. When the normality of distribution of within subject errors was not assumed, the fit statistics hit rate decreased of 2 points as average.

Table 3. Hit rate of fit statistics for searching covariance structures (A: AR(1) structure, M: MA(1) structure, N: ARMA(1,1) structure): Mean Values (%)

Simulation	AIC			BIC			CAIC			HQIC			AICC		
conditions	A	M	N	A	M	N	A	M	N	A	M	N	A	M	N
N with 50 subjects and 5 time points	40	75	6	22	57	0	16	56	0	30	69	0	39	75	5
N with 50 subjects and 8 time points	16	58	82	55	56	18	56	56	10	37	58	53	17	58	81
N with 100 subjects and 5 time points	51	81	22	24	85	0	21	84	0	44	87	8	50	82	21
N with 100 subjects and 8 time points	26	89	78	51	91	24	52	88	17	38	89	58	27	89	78
N with 150 subjects and 5 time points	25	82	40	13	78	1	10	72	1	20	85	9	25	82	37
N with 150 subjects and 8 time points	7	78	92	26	85	43	27	82	36	15	95	69	7	78	92
N with 200 subjects and 5 time points	36	63	6	24	84	0	20	81	0	31	78	2	36	63	6
N with 200 subjects and 8 time points	29	42	64	47	88	15	47	92	11	44	68	40	29	45	64
C with 50 subjects and 5 time points	31	66	7	15	56	0	12	51	0	24	65	1	30	66	5
C with 50 subjects and 8 time points	19	62	76	47	57	20	49	55	10	37	60	45	21	62	75
C with 100 subjects and 5 time points	46	86	18	24	81	0	19	78	0	42	87	5	46	86	17
C with 100 subjects and 8 time points	27	85	68	42	80	18	44	76	12	37	87	42	27	85	67
C with 150 subjects and 5 time points	26	77	28	13	74	2	13	69	2	21	82	9	26	77	27
C with 150 subjects and 8 time points	9	65	88	28	72	35	29	69	25	15	76	62	9	65	88
C with 200 subjects and 5 time points	28	58	11	16	81	0	14	78	0	24	78	5	29	58	11
C with 200 subjects and 8 time points	37	38	54	52	89	14	50	92	10	45	67	32	37	39	54
Growth parameter (β_1=0.05)	28	69	46	31	76	12	30	74	8	32	77	27	28	70	46
Growth parameter (β_1=0.16)	28	69	46	31	76	12	30	74	8	31	77	27	28	69	45
G matrix (τ_{00}=0.1 and τ_{11}=0.05)	40	78	48	53	90	15	52	90	12	49	87	30	40	78	47
G matrix (τ_{00}=0.2 and τ_{11}=0.10)	17	60	45	10	62	9	8	57	5	14	67	25	17	61	44

N: normal distribution of within subject error, C: Chi square (1) distribution of within subject error. The standard deviation of hit rate of fit statistics for searching AR(1), MA(1) and ARMA(1,1) structures ranged respectively from 2.37 to 42.09 %, 0.39 to 34.22 % and 0.00 to 37.23 %.

About MA(1) covariance structure, it appears from the results of ANOVA that only the G matrix significantly affected all hit rates (not presented). Moreover, BIC and CAIC was moderated by the sample size. From the mean values of fit statistics performance (Table 3), BIC and CAIC were able to correctly classify the covariance structure 57 % and 54 % of the time (lowest values) with 50 in-dividuals respectively and 86 % and 85 % of the time (highest value) with 200 individuals respectively. With reregards to AIC and AICC, the lowest hit rates were found for 200 individuals and the highest hit rates were found for 150 individuals. 63 % (50 individuals) was the lowest HQIC hit rate and 88 % (100 individuals) was the highest HQIC hit rate. Regarding the G matrix, the highest values for all hit rates were found for small size of G matrix. When the normality of within subject errors was not assumed, the fit statistics hit rate decreased of 4 points as average.

When ARMA(1,1) covariance structure is considered, the results of ANOVA applied on fit statistics hit rates to inves-tigate the effect of design factors indicate that only meas-urement periods moderated the fit statistics hit rate (not presented). The inspection of mean values of fit statistics performance according to the measurement periods (Table 3) reveals that the highest values for all hit rates were found for 8 measurement periods (76 % of the time for AIC, 23 % of the time for BIC, 16 % of the time for CAIC, 50 % of the time for HQIC and 75 % of the time for AICC). When the normality of distribution of within subject errors was not assumed, the fit statistics hit rate decreased of 4 points as average.

3.2. Consequence of misspecification in within-subject variance-covariance structure

3.2.1. Relative bias

Summary statistics for the mean relative bias (MRB) of the fixed effects are putted in Table 4. It shows that the mean and median for β_0 and β_1=0.16 were very close to zero whereas the second slope term (i.e. β_1=0.05) had much more variation. These terms also have a few small relative bias statistics shown by the small minimum and maximum values. Therefore, on average, the relative bias was kept under control for all of the fixed effects, but can become a problem for the slope term β_1=0.05. On average, the random components tended to be biased and there was large variation in the RB statistics for each term.

Table 4. Summary Statistics for Relative Bias of Fixed Effects and Random Components

Term	Mean	Var	Med	Min	Max
β_0	-0.0055	0.0020	-0.0031	-0.1481	0.1239
β_1=0.05	0.1324	0.8734	0.4356	-2.3555	1.8066
β_1=0.16	-0.0238	0.1832	-0.0063	-0.7660	1.0912
τ_{00}=0.10	11.1703	15.6449	11.6894	2.3226	16.9867
τ_{00}=0.20	4.3273	2.7114	4.2420	-0.2352	7.3008
τ_{11}=0.05	0.5041	0.4394	0.3903	-0.5610	2.0901
τ_{11}=0.10	-0.1931	0.1993	-0.3513	-0.7406	0.8798
τ_{01}	-4.5152	6.3786	-4.0672	-10.9275	-0.9738
$\sigma^2\epsilon$	-0.3336	0.0188	-0.3521	-0.5904	-0.0348

Var: variance, Med: median, Min: minimum, Max: maximum.

To better explore the variability shown in Table 4, boxplots of the fixed effects and random components are shown in Figure 1. It reveals a large amount of variability in β_1=0.05

compared to others fixed effects. Moreover, we notice a large amount of variability in τ_{00} and τ_{01}. The interquartile range, depicted by the box in the boxplots, was much larger for these terms indicating that the RB has a wider range of plausible values.

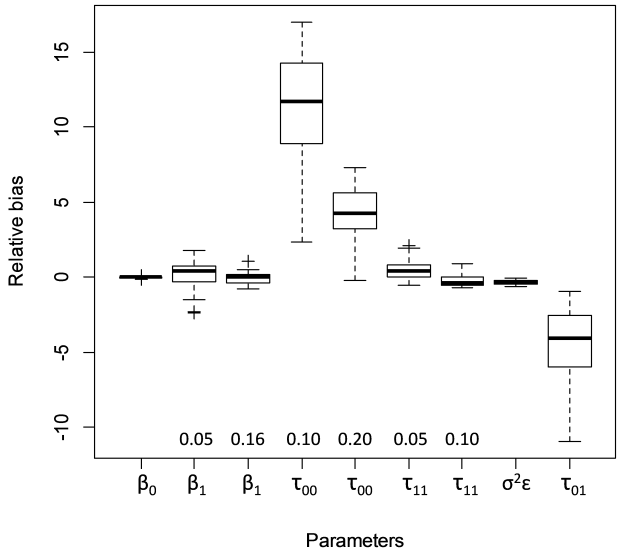


Fig. 1. Boxplots showing the relative bias of the fixed effects and random components

3.2.2. Relative Error

Whichever the fixed effect considered, on average the relative error was kept under control without important difference between covariance structures (Table 5). The perfomance of covariance structures was the same among the considered factors for model intercept variance (τ_{00}), covariance between β_0 and β_1 (τ_{01}) and variance of within subject errors ($\sigma^2\epsilon$), except for the model slope (τ_{11}). Indeed, ARMA(1,1) covariance structure records the best performance followed by AR(1) covariance structure. Regarding the model slope (τ_{11}), the performance of the fitted covariance structures depends on the characteristics of the considered sample.

Boxplots of the MRE of the fitted covariance structures (Figure 2) shows almost the same performance of the fitted covariance structures for the fixed effects. Regarding the random components, the performance of fitted covariance structures becomes greatly different. ARMA(1,1) covariance structure followed by AR(1) covariance structure, with however, important gap between both covariance structures presented the best performance for the model intercept variance, covariance between the model intercept and the model slope and variance of within subject errors. On the contrary, the lowest performances occurred when the covariance structure was underspecified i.e. ID structure fit to an AR(1) structure. Moreover, the dispersions of the MRE confirm the previous observations.

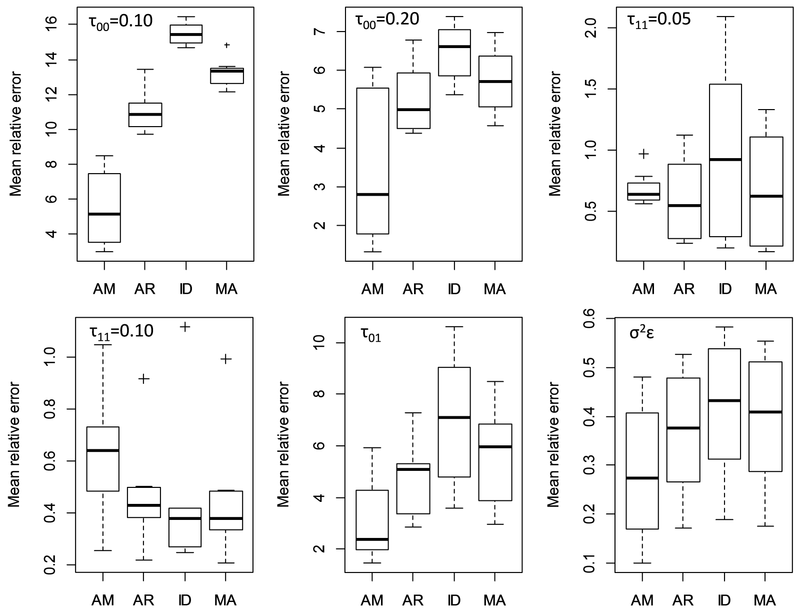
Table 5. Median ranks of the relative error of random components according to the considered factors (I: independence structure, A: AR(1) structure, M: MA(1) structure, N: ARMA(1,1) structure)

Simulation	τ_{00}=0.10				τ_{00}=0.20				τ_{11}=0.05				τ_{11}=0.10				τ_{01}				$\sigma^2\epsilon$			
conditions	I	A	M	N	I	A	M	N	I	A	M	N	I	A	M	N	I	A	M	N	I	A	M	N
Overall	4	2	3	1	4	2	3	1	3.5	2	2	2.5	1.5	3	2	4	4	2	3	1	4	2	3	1
50 subjects and 5 time points	4	2	3	1	4	2	3	1	4	2	3	1	3	2	1.5	3.5	4	2	3	1	4	2	3	1
50 subjects and 8 time points	4	2	3	1	4	2	3	1	3	2	1	4	1	3	2	4	4	2	3	1	4	2	3	1
100 subjects and 5 time points	4	2	3	1	4	2	3	1	4	2	3	1	3	1.5	1.5	4	4	2	3	1	4	2	3	1
100 subjects and 8 time points	4	2	3	1	4	2	3	1	2	3	1	4	1	3	2	4	4	2	3	1	4	2	3	1
150 subjects and 5 time points	4	2	3	1	4	2	3	1	4	2	3	1	3	2	1.5	3.5	4	2	3	1	4	2	3	1
150 subjects and 8 time points	4	2	3	1	4	2	3	1	3	2	1	4	1	3	2	4	4	2	3	1	4	2	3	1
200 subjects and 5 time points	4	2	3	1	4	2	3	1	4	2	3	1	3	1.5	1.5	4	4	2	3	1	4	2	3	1
200 subjects and 8 time points	4	2	3	1	4	2	3	1	2	3	1	4	1	3	2	4	4	2	3	1	4	2	3	1
Growth parameter (β_1=0.05)	4	2	3	1	4	2	3	1	3.5	2	2	2.5	2	2	2	4	4	2	3	1	4	2	3	1
Growth parameter (β_1=0.16)	4	2	3	1	4	2	3	1	3.5	2	2	2.5	1.5	3	1.5	4	4	2	3	1	4	2	3	1
G matrix (τ_{00}=0.1 and τ_{11}=0.05)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4	2	3	1	4	2	3	1
G matrix (τ_{00}=0.2 and τ_{11}=0.10)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4	2	3	1	4	2	3	1

4. Discussion

This study was firstly intended to assess the performance of five fit statistics on searching for the correct within-subject covariance structure in longitudinal data analysis, and then to examine the effect of misspecification in the within-subject variance covariance structure on the model estimations considering some characteristics of the sample study namely the sample size, the measurement periods, the magnitude of growth parameter, the size of G matrices, the covariance structure of the within-subject error and the distribution of the within-subject error.

The results suggested that the individual average performance ranked HQIC, BIC and CAIC for searching AR(1) and MA(1) covariance structures as the top three fit statistics. In particular, the AIC and AICC correct identification for ARMA(1,1) outperformed all the other fit statistics classification for ARMA(1,1). AIC and AICC hit rates were higher than HQIC hit rate, BIC hit rate or CAIC hit rate. It is known that the more parameters there are in the model, the better the fit (Crawley (2013)). However, despite of AIC penalizes complex model due to many parameters being estimated and promotes parsimonious models (Crawley (2013), Lee (2010)), it had the best hit rate for ARMA(1,1) covariance structure. The obtained performance of fit statistics in this study contrasts with those found by Ferron et al. (2002) which were overall higher ranging from 66 % for BIC to 79 % for AIC and by Lee (2010) [66 % for BIC and 62 % for AIC].
As pointing out in previous studies (Keselman *et al.* (1998), Wolfinger (1993)), some fit index did not perform well on searching for the optimal covariance structure under the general mixed model framework. Indeed, according to Keselman *et al.* (1998), AIC and BIC performance ranged from 47 % to 35 % respectively. It has also been revealed that measurement periods (Lee (2010)) current study) and the size of G matrix (current study) played a major role in the variability of the hit rate in these fit statistics. For these fit statistics, small time points and high size of G matrix had implied lowest performance.



ID: independence structure, AR: AR(1) structure, MA: MA(1) structure, AM: ARMA(1,1) structure

Fig. 2. Boxplots of the mean relative error per random components and fitted covariance structure

Moreover, the sample size moderated BIC and CAIC hit rate for searching MA(1) covariance structure.

The study also examined the effect of misspecification in the within-subject variance covariance structure on the model estimations. Our findings suggested that across all models, the estimates for the fixed effects mainly for β_0 and β_1=0.16 were almost unbiased (relative bias $<$ 0.05). This finding was consistent with previous research when the within-subject error structure is misspecified (Ferron *et al.* (2002); Brandon (2013)). On the contrary, the random components were clearly biased with misspecification of the model. The most biased parameters were the variance of model intercept and the covariance between the model intercept and the model slope. Indeed, the aim of using linear mixed effects models to analyse longitudinal data is to carry out the difference between repeated measures within individuals and the difference between subjects. In spite of knowing that the variations among individuals in the interested variable over time is reflected in the covariance matrix of the within-subject residuals, (Hedeker and Gibbons (2006)), the random effects are those that summarize the between subjects difference. Therefore, the misspecification of covariance structure implies that LMEM fails to well model the between subjects difference. Moreover, the same conclusions were drawn from the relative errors showing the worst performance of model misspecification, mainly the under-specification or generally misspecification in the covariance structure of within subject errors. The largest amounts of errors were found when the fitted structure was underspecified as independent

(ID).

5. Conclusion

This study firstly examined the performance of five fit statistics in selecting the more suitable within-subject variance covariance matrix in longitudinal data analysis under linear mixed effects models framework. The averaged overall hit rates for AIC, BIC, CAIC, HQIC and AICC were below 50 %. The worst concern as to these fit statistics was that their stability in searching for the optimal covariance structure aggravated as the target covariance structure became more complex. Based on the overall and steady performance of HQIC and BIC for generated variance covariance matrix, we concluded these criteria had better ability in assessing of optimal within-subject covariance structure. The study results also showed that the fixed effects on average were unbiased and the size of G matrix, the covariance structure and the measurement periods were the design factors that explain the variations in the mean relative bias for the fixed effects. However, there was evidence of bias in the random components and some simulation conditions, namely measurement periods and covariance structure did explain significant variation in the average relative bias. Regarding the relative errors, the random components tend to be overestimated when covariance structure is underspecified.

References

Akaike, H.(1974). A new look at the statistical model identification. IEEE Trans. *Automatic Cont.*, 19: 716-723. Doi : 10.1109/TAC.1974.1100705

AL-Marshadi, A.H.(2014). Selecting the covariance structure in mixed model using statistical methods calibration. *Journal of Mathematics and Statistics*, 10(3), 309-315. Doi : 10.3844/jmssp.2014.309.315

Bozdogan, H.(1987). Model selection and Akaike's infor-mation criterion (AIC): the general theory and its analytical extensions. *Psychometrika*, 52, 345-370. Doi : 10.1007/BF02294361

Brandon, L.(2013). *Misspecification of the covariance matrix in the linear mixed model: A Monte Carlo simulation*. PhD thesis, University of Minnesota, MN, USA.

Crawley, M.J.(2013). *The R book*. The second edition. John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, United Kingdom.

Ferron, J., Dailey, R., and Yi, Q.(2002). Effects of misspecifying the first-level error structure in two-level models of change. *Multivariate Behavioral Research*, 37(3), 379–403. Doi : 10.1207/S15327906MBR37034

Gibbons, R.D., Hedeker, D., and DuToit, S.(1979). *Advances in Analysis of Longitudinal Data. Reviews in Advance.*

Hannan, E.J., and Quinn, B.G.(1979). The determination of the order of an autoregression. *Multivariate Behavioral Journal of Royal Statistical Society Series B*, 41(2), 190-195.

Hedeker, D., and Mermelstein, R.J.(2007). Mixed-effects regression models with heterogeneous variance: Analyzing ecological momentary assessment (EMA) data of smoking. *In : Little, T.D., Bovaird, J.A., Card, N.A., (Eds.), Modeling contextual effects in longitudinal studies*, p. 183-206. Mahwah, NJ: LEA.

Hedeker, D.R., and Gibbons, R.D.(2006). *Longitudinal data anal-ysis.* Malden, MA: Wiley-Interscience.

Hurvich, C.M., and Tsai, C.L.(1989). Regression and time series model selection in small samples. *Biometrika*, 76, 297-307. Doi : 10.1093/biomet/76.2.297

Keselman, H., Algina, J., Kowalchuk, R., and Wolfinger, R.(1998). A comparison of two approaches for selecting covariance structures in the analysis of repeated measurements. *Communications in Statistics-Simulation and computation*, 27 (3), 591-604. Doi : 10.1080/03610919808813497

Lee, Y-H.(2010). *Longitudinal data analysis using multilevel linear Modeling (MLM): Fitting an optimal vari-ance-covariance structure*. PhD thesis, Texas A&M Univer-sity, USA.

Maas, C.J.M., and Hox, J.J.(2004). The influence of violations of assumptions on multilevel parameter estimates and their standard errors. *Computational Statistics and Data Analysis*, 46, 427-440. Doi : 10.1016/j.csda.2003.08.006

McCulloch, C.E.(2003). *Generalized Linear Mixed Models.* 1st Edn., Illustrated, IMS, Beachwood, ISBN-10: 0940600544.

Micceri, T.(1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin*, 105, 156-166. Doi : 10.1037/0033-2909.105.1.156

Schwarz, G.E.(1978). Estimating the dimension of a model. *Annals of Statistics*, 6 (2): 461–464. DOI : 10.1214/aos/1176344136

Snijders, T.A.B., and Bosker, R.(1999). *Multilevel Analysis. An Introduction to Basic and Advances Multilevel Modeling.* Sage. Thousand Oaks, CA.

Steele, F.(2008). Multilevel models for longitudinal data. *Journal of the Royal Statistical Society: series A (statistics in society)*, 171 (1), 5-19. Doi : 10.1111/j.1467-985X.2007.00509.x/full

Wolfinger, R.(1993). Covariance structure selection in general mixed models. *Communications in Statistics: Simulation and Computation*, 22(4), 1079–1106. Doi : 10.1080/03610919308813143

Yanosky II, D.J.(2007). *Comparability of covariance struc-tures and accuracy of information criteria in mixed model methods for longitudinal data analysis.* PhD thesis, Univer-sity of Georgia, Athens, Georgia, USA.