Available online at http://scik.org
Commun. Math. Biol. Neurosci. 2020, 2020:27
https://doi.org/10.28919/cmbn/4610
ISSN: 2052-2541

LATENT FACTOR LINEAR MIXED MODEL (LFLMM) FOR MODELLING FLANDERS DATA

YENNI ANGRAINI[1,*], KHAIRIL ANWAR NOTODIPUTRO[1], ASEP SAEFUDDIN[1], TONI TOHARUDIN[2]

[1]Department of Statistics, IPB University, Bogor, Indonesia

[2]Department of Statistics, University of Padjadjaran, Bandung, Indonesia

Abstract: Latent factor linear mixed model (LFLMM) is a method that generally used for the analysis of change in high- dimensional longitudinal data. The LFLMM framework works under the linear mixed model framework. Analysis of change from several latent variables such as Individualism (I), Nationalism (N), Ethnocentrism (E), and Authoritarianism (A) in Flanders, Belgium, is interesting as Belgium is feared to fall apart as a nation. Two main research questions in the Flanders case are how Individualism (I), Nationalism (N), Ethnocentrism (E), and Authoritarianism (A) develop over time and whether there exist association between the Individualism (I), Nationalism (N), Ethnocentrism (E), and Authoritarianism (A) developments. Although these latent variables have been the subject of several studies in Flanders, an analysis of all four concepts using Latent factor linear mixed model has not been performed. Hence, it is the interest of this paper to discuss such model using the Flanders data. Two stages of modelling have been carried out. The first stage involved modelling Individualism (I), Nationalism (N), and Ethnocentrism (E) and in the next stage Authoritarianism (A) was added to the model. The results showed that I, N, and A increased over time while E decreased over time. The correlation of random effects in LFLMM suggest several interesting findings, including the positive correlation between E and A; I and E; and also between I and A.

Keywords: latent factor linier mixed model; longitudinal data analysis; Flanders-Belgium.

2010 AMS Subject Classification: 62P25, 62H99.

---

[*]Corresponding author
E-mail address: y_angraini@apps.ipb.ac.id
Received April 8, 2020

1. INTRODUCTION

A joint modelling strategy is needed to answer some research questions which are of interest in assessing the relation between covariates and all outcomes simultaneously, in studying how the association between the various outcomes evolves over time, or in investigating the association between the evolutions of all outcomes [1]. A number of approaches to joint modelling or simultaneous modelling have been proposed to handle multivariate or multiple outcomes in longitudinal data. There are two ways to solve this (i) to use one or more latent variables for the outcome dimension, which is to reduce the dimensions of the multivariate/multiple outcomes, (ii) to use latent variables for the time dimension, i.e. to assume that the repeated measurements of a particular outcome (one frequency at either left or right side) are reflecting a latent evolution for that outcome ([2]–[5]). For example, in social sciences, such as psychometrics, structural equation model (SEM) are often used to model longitudinal data [6], [7]. According to Bollen and Curran [8], this approach has several appealing modelling abilities and can be used for multiple outcomes in longitudinal data or high-dimensional data. Under the SEM and continuous time models frame work ([9]–[11]), Voelkle et al. [12] proposed the continuous time in SEM (CT-SEM) model to avoid some issues associated with the autoregressive and cross-lagged models. The main challenge in most latent variable models that it is complicated to use especially for non-statistical background users since the marginal likelihood function is not straightforward, involving the integration on latent variables[13], [14], [15].

An et al. [16] proposed another approach to handle high-dimensional data in outcomes. This model was developed to address the issue of interrelated trends among latent variables that can only be addressed by modelling the latent variables jointly. It reduced the high-dimensional responses to low-dimensional latent factors by the factor analysis model [17], [18], and then used the multivariate linear mixed model to study the longitudinal trends of these latent factors, where the estimates have been done using the EM algorithm. Two main research questions that can be answered by latent factor linear mixed model (LFLMM) are (i) how the latent variables develop over time and (ii) whether there is any association between the latent variables developments. They can be answered by the fixed slope and the correlation matrix of random effects. Kondaurova et al. [19] applied LFLMM to study the affective properties of infant-directed speech influence the attention of infants with normal hearing to speech sounds. By using LFLMM, this study examined (a) whether the perception of affective and directive qualities of Infant-Directed-Speech (IDS) depended on the hearing status of the infant and (b) whether the perception of affective and directive qualities of IDS changed over the period of the three testing sessions in low-pass-filtered (LPF) speech in three groups (Mothers of infants with hearing impairment, mothers of infants with normal hearing matched by age and mothers of infants with normal hearing experience).

CT-SEM and LFLMM work under different framework. The framework for CT-SEM is SEM where the assumptions and limitations of SEM apply to the CT-SEM while LFLMM works under the linear mixed model framework. These two models are generally the methods used for the analysis of change in high-dimensional longitudinal data ([12], [16]).

Analysis of the relationships among Individualism (I), Nationalism (N), Ethnocentrism (E), and Authoritarianism (A) in Flanders, Belgium, have been the topic of several studies. Toharudin et. al [20] showed that I and E are connected in a moderately strong feedback relationship with the effect from I towards E somewhat stronger than that in the opposite direction. Furthermore, both I and E have small effects on N. By adding the variable Authoritarianism (A), Angraini et al. [21] found reciprocal effects between A and E and between E and I as well as a unidirectional effect from A on I. The first paper used the Bergstrom's approximate discrete model while the second paper used the continuous-time SEM (CT-SEM) approach.

Analysis of change from several latent variables such as Individualism (I), Nationalism (N), Ethnocentrism (E), and Authoritarianism (A) in Flanders, Belgium, is interesting to be investigated further as Belgium is feared to fall apart as a nation. Identification with Belgium as a nation has always been relatively weak, even before the start of the federalization process. As explained in the previous paragraph, simultaneous analysis of the four latent variables so far has only used Bergstrom's approximate discrete model and the continuous time SEM (CT-SEM) model. While there has never been a researcher who used the LFLMM approach. The objective of this paper is to study the analysis of Flanders data using LFLMM to answer the research question: how the Individualism (I), Nationalism (N), Ethnocentrism (E), and Authoritarianism (A) in Flanders develop over time and whether there is any association between the Individualism (I), Nationalism (N), Ethnocentrism (E), and Authoritarianism (A) in Flanders developments. We carried out two stages of modelling, first involving only Individualism (I), Nationalism (N), Ethnocentrism (E) for both methods then adding Authoritarianism (A) to the next stage. This was done to see the consistency of the changes in the three latent variables with the inclusion of Authoritarianism (A).

The paper is organized as follows. Section 2 presents the Latent Factor Linear Mixed Model. In section 3, we discuss sample and variables. Section 4 discusses the results. Finally, conclusion is presented in section 5.

2. LATENT FACTORS LINEAR MIXED MODEL

Latent Factors Linear Mixed Model is a method that have been proposed by An et al. [16] to handle multiple outcomes in longitudinal data. The modelling framework was used in LFLMM similar to that of Roy and Lin [22] proposed. The general idea is to use a factor-analytic for reducing the dimension of response vector and next to use standard longitudinal models, linear mixed model, for analysing the longitudinal trends of low-dimension of response vector. In matrix notation, the specification models of LFLMM present as

$$\boldsymbol{Y}_i = (\boldsymbol{I}_{T_i} \otimes \boldsymbol{\Lambda})\boldsymbol{\eta}_i + \boldsymbol{\epsilon}_i \quad (1)$$

$$\boldsymbol{\eta}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{a}_i + \boldsymbol{\varepsilon}_i \quad (2)$$

where

$$\boldsymbol{Y}_i = (y'_{i1}, \dots, y'_{iT_i})'_{[J\times T_i,1]}$$

$$\boldsymbol{\eta}_i = (\eta'_{i1}, \dots, \eta'_{iT_i})'_{[d\times T_i,1]}$$

$$\boldsymbol{\epsilon}_i = (\boldsymbol{\epsilon}'_{i1}, \dots, \boldsymbol{\epsilon}'_{iT_i})'_{[J\times T_i,1]}$$

$$\boldsymbol{\Lambda}_{[J\times d]} = \begin{pmatrix} \lambda'_1 \\ \vdots \\ \lambda'_J \end{pmatrix}, \ \boldsymbol{X}_i = \begin{pmatrix} \boldsymbol{x}_{i1} \\ \vdots \\ \boldsymbol{x}_{iT_i} \end{pmatrix}_{[d\times T_i, p\times d]}, \boldsymbol{Z}_i = \begin{pmatrix} z_{i1} \\ \vdots \\ z_{iT_i} \end{pmatrix}_{[d\times T_i, q\times d]}$$

$$\boldsymbol{\beta} = (\boldsymbol{\beta}^{1'}, \dots, \boldsymbol{\beta}^{d'})'_{[p\times d,1]}$$

$$\boldsymbol{a}_i = (\boldsymbol{a}_i^{1'}, \dots, \boldsymbol{a}_i^{d'})'_{[q\times d,1]} \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}_a)$$

$$\boldsymbol{\varepsilon}_i = (\boldsymbol{\varepsilon}'_{i1}, \dots, \boldsymbol{\varepsilon}'_{iT_i})'_{[d\times T_i,1]}, \ \ \boldsymbol{\varepsilon}_{it} \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}_\varepsilon)$$

The parameter estimation for both models is done by EM algorithm approach [23]. According to An et al. [16], the advantages of using EM algorithms in this model are easy to implement because a closed-form solution is available for both the E-step and M-step steps in the EM algorithm. The log-likelihood equation is

$$lnL = \sum_{i=1}^{N}\left[logP(\boldsymbol{Y}_i|\boldsymbol{\eta}_i, \boldsymbol{\Lambda}, \boldsymbol{\tau^2}) + logP(\boldsymbol{\eta}_i|\boldsymbol{X}_i, \boldsymbol{Z}_i, \boldsymbol{a}_i, \boldsymbol{\beta}, \boldsymbol{\Sigma}_\varepsilon) + logP(\boldsymbol{a}_i|\boldsymbol{\Sigma}_a)\right] \quad (3)$$

The following is the estimator obtained after maximizing the likelihood log function of the model used in M-step of the EM algorithm,

$$\widehat{\boldsymbol{\Lambda}}' = \left[\sum_{i=1}^{m}\sum_{t=1}^{n_i}\boldsymbol{\eta}_{it}\boldsymbol{\eta}'_{it}\right]^{-1}\left[\sum_{i=1}^{m}\sum_{t=1}^{n_i}\boldsymbol{\eta}_{it}\boldsymbol{y}'_{it}\right] \quad (4)$$

$$\hat{\boldsymbol{\tau}}_j{}^2 = \frac{1}{n}diag\left[\sum_{i=1}^{m}\sum_{t=1}^{n_i}\boldsymbol{y}_{it}\boldsymbol{y}'_{it} - 2\sum_{i=1}^{m}\sum_{t=1}^{n_i}\boldsymbol{y}_{it}\boldsymbol{\eta}'_{it}\widehat{\boldsymbol{\Lambda}}' + \widehat{\boldsymbol{\Lambda}}\sum_{i=1}^{m}\sum_{t=1}^{n_i}\boldsymbol{\eta}_{it}\boldsymbol{\eta}'_{it}\widehat{\boldsymbol{\Lambda}}'\right] \quad (5)$$

$$\widehat{\boldsymbol{\Sigma}}_a = \frac{1}{m}\sum_{i=1}^{m}\boldsymbol{a}_i\boldsymbol{a}'_i \quad (6)$$

The predictors of $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}_\varepsilon$ can be obtained through the iteration process. Let $\boldsymbol{\Sigma}_\varepsilon$ be determined then the predictor for β is

$$\widehat{\boldsymbol{\beta}} = \left[\sum_{i=1}^{m}\sum_{t=1}^{n_i}\boldsymbol{x}'_{it}\boldsymbol{\Sigma}_\varepsilon^{-1}\boldsymbol{x}_{it}\right]^{-1}\left[\sum_{i=1}^{m}\sum_{t=1}^{n_i}\boldsymbol{x}'_{it}\boldsymbol{\Sigma}_\varepsilon^{-1}(\boldsymbol{\eta}_{it} - z_{it}\boldsymbol{a}_i)\right] \quad (7)$$

and let $\boldsymbol{\beta}$ be determined then

$$\widehat{\boldsymbol{\Sigma}}_\varepsilon = \frac{1}{n}\sum_{i=1}^{m}\sum_{t=1}^{n_i}\left[\boldsymbol{\eta}_{it} - \boldsymbol{x}_{it}\widehat{\boldsymbol{\beta}} - z_{it}\boldsymbol{a}_i\right]\left[\boldsymbol{\eta}_{it} - \boldsymbol{x}_{it}\widehat{\boldsymbol{\beta}} - z_{it}\boldsymbol{a}_i\right]' \quad (8)$$

As for E-step, An et al. [16] performs calculations for conditional expect values $\boldsymbol{\eta}_{it}, \boldsymbol{\eta}_{it}\boldsymbol{\eta}'_{it}, \boldsymbol{a}_i, \boldsymbol{a}_i\boldsymbol{a}'_i$ and $\boldsymbol{\eta}_{it}\boldsymbol{a}'_i$. To derive the conditional expectation value of the five statistics, the joint distribution of $\boldsymbol{Y}_i, \boldsymbol{\eta}_i$ and $\boldsymbol{a}_i$ is as follows:

$$\begin{pmatrix}\boldsymbol{Y}_i\\ \boldsymbol{\eta}_i\\ \boldsymbol{a}_i\end{pmatrix} \sim N\left(\begin{pmatrix}\boldsymbol{I}_{n_i}\otimes\boldsymbol{\Lambda}\boldsymbol{x}_i\boldsymbol{\beta}\\ \boldsymbol{x}_i\boldsymbol{\beta}\\ \boldsymbol{0}\end{pmatrix}, \begin{pmatrix}\boldsymbol{\Sigma}_{y_i} & \boldsymbol{\Sigma}_{y_i\eta_i} & \boldsymbol{\Sigma}_{y_ia_i}\\ \boldsymbol{\Sigma}_{\eta_iy_i} & \boldsymbol{\Sigma}_{\eta_i} & \boldsymbol{\Sigma}_{\eta_ia_i}\\ \boldsymbol{\Sigma}_{a_iy_i} & \boldsymbol{\Sigma}_{a_i\eta_i} & \boldsymbol{\Sigma}_{a_i}\end{pmatrix}\right)$$

where

$$\boldsymbol{\Sigma}_{y_i} = (\boldsymbol{I}_{n_i}\otimes\boldsymbol{\Lambda})\boldsymbol{\Sigma}_{\eta_i}(\boldsymbol{I}_{n_i}\otimes\boldsymbol{\Lambda})' + \boldsymbol{I}_{n_i}\otimes diag(\boldsymbol{\tau}_1^2, \dots, \boldsymbol{\tau}_J^2) \quad (9)$$

$$\boldsymbol{\Sigma}_{\eta_i} = \boldsymbol{z}_i\boldsymbol{\Sigma}_a\boldsymbol{z}'_i + \boldsymbol{I}_{n_i}\otimes\boldsymbol{\Sigma}_\varepsilon \quad (10)$$

$$\boldsymbol{\Sigma}_{a_i} = \boldsymbol{\Sigma}_a \quad (11)$$

$$\boldsymbol{\Sigma}_{y_i\eta_i} = (\boldsymbol{I}_{n_i}\otimes\boldsymbol{\Lambda})\boldsymbol{\Sigma}_{\eta_i} \quad (12)$$

$$\boldsymbol{\Sigma}_{y_ia_i} = (\boldsymbol{I}_{n_i}\otimes\boldsymbol{\Lambda})(\boldsymbol{z}_i\boldsymbol{\Sigma}_a) \quad (13)$$

$$\boldsymbol{\Sigma}_{\eta_ia_i} = \boldsymbol{z}_i\boldsymbol{\Sigma}_a \quad (14)$$

The conditional distribution of $\boldsymbol{\eta}_i$ and $\boldsymbol{a}_i$ is the normal disribution with the mean value and the matrix of the uniform variety as follows,

$$\boldsymbol{\mu}_{\eta_i|y_i} = \boldsymbol{x}_i\boldsymbol{\beta} + \boldsymbol{\Sigma}_{\eta_i y_i}\boldsymbol{\Sigma}_{y_i}^{-1}\left(\boldsymbol{y}_i - \boldsymbol{I}_{n_i} \otimes \boldsymbol{\Lambda}\boldsymbol{x}_i\boldsymbol{\beta}\right) \tag{15}$$

$$\boldsymbol{\Sigma}_{\eta_i|y_i} = \boldsymbol{\Sigma}_{\eta_i} - \boldsymbol{\Sigma}_{\eta_i y_i}\boldsymbol{\Sigma}_{y_i}^{-1}\boldsymbol{\Sigma}_{y_i\eta_i} \tag{16}$$

$$\boldsymbol{\mu}_{a_i|y_i} = \boldsymbol{\Sigma}_{a_i y_i}\boldsymbol{\Sigma}_{y_i}^{-1}\left(\boldsymbol{y}_i - \boldsymbol{I}_{n_i} \otimes \boldsymbol{\Lambda}\boldsymbol{x}_i\boldsymbol{\beta}\right) \tag{17}$$

$$\boldsymbol{\Sigma}_{a_i|y_i} = \boldsymbol{\Sigma}_{a_i} - \boldsymbol{\Sigma}_{a_i y_i}\boldsymbol{\Sigma}_{y_i}^{-1}\boldsymbol{\Sigma}_{y_i a_i} \tag{18}$$

The joint conditional distribution of $\boldsymbol{\eta}_i$ and $\boldsymbol{a}_i$ follows a normal distribution with a matrix of varying degrees,

$$\boldsymbol{\Sigma}_{\eta_i, a_i|y_i} = \begin{pmatrix} \boldsymbol{\Sigma}_{\eta_i} & \boldsymbol{\Sigma}_{\eta_i a_i} \\ \boldsymbol{\Sigma}_{a_i\eta_i} & \boldsymbol{\Sigma}_{a_i} \end{pmatrix} - \begin{pmatrix} \boldsymbol{\Sigma}_{\eta_i y_i} \\ \boldsymbol{\Sigma}_{a_i y_i} \end{pmatrix} \boldsymbol{\Sigma}_{y_i}^{-1} \left(\boldsymbol{\Sigma}_{y_i\eta_i} \quad \boldsymbol{\Sigma}_{y_i a_i}\right) \tag{19}$$

and the conditional uniform matrix of $\boldsymbol{\eta}_i$ and $\boldsymbol{a}_i$ are as follows,

$$\boldsymbol{\Sigma}_{\eta_i\, a_i|y_i} = \boldsymbol{\Sigma}_{\eta_i a_i} - \boldsymbol{\Sigma}_{\eta_i y_i}\boldsymbol{\Sigma}_{y_\text{i}}^{-1}\boldsymbol{\Sigma}_{y_i a_i} \tag{20}$$

thus

$$E(\boldsymbol{\eta}_{it}|\boldsymbol{y}_i) = \boldsymbol{\mu}_{\eta_i|y_i}[1 + (t-1) * d : t * d] \tag{21}$$

$$E(\boldsymbol{\eta}_{it}\boldsymbol{\eta}'_{it}|\boldsymbol{y}_i) = E(\boldsymbol{\eta}_{it}|\boldsymbol{y}_i)E(\boldsymbol{\eta}_{it}|\boldsymbol{y}_i)' + \boldsymbol{\Sigma}_{\eta_i|y_i}[1 + (t-1) * d : t * d, 1 + (t-1) * d : t * d] \tag{22}$$

$$E(\boldsymbol{a}_i|\boldsymbol{y}_i) = \boldsymbol{\mu}_{a_i|y_i} \tag{23}$$

$$E(\boldsymbol{a}_i\boldsymbol{a}'_i|\boldsymbol{y}_i) = E(\boldsymbol{a}_i|\boldsymbol{y}_i)E(\boldsymbol{a}_i|\boldsymbol{y}_i)' + \boldsymbol{\Sigma}_{a_i|y_i} \tag{24}$$

$$E(\boldsymbol{\eta}_{it}\boldsymbol{a}'_i) = E(\boldsymbol{\eta}_{it}|\boldsymbol{y}_i)E(\boldsymbol{a}_i|\boldsymbol{y}_i)' + \boldsymbol{\Sigma}_{\eta_i\, a_i|y_i}[1 + (t-1) * d : t * d] \tag{25}$$

3. SAMPLE AND VARIABLES

The General Election Study (Interuniversitair Steunpunt Politieke-opinieonderzoek) in Belgium was designed to include a representative sample of the target population under the Belgian electorate. The sample as presented in Toharudin et al. [20] had two type respondents, the Flemish respondents and Dutch-speaking respondents of the Brussels-Capital Region. We called the data set as the Flemish data set. It consists of 1274 respondents, who have been interviewed three times i.e 1991, 1995 and 1999 ([24]–[26]). For more detailed information about the sample, see [20] and [21].

There are four latent variables measured. Individualism is measured by four 5-point-scale items: Everybody has to take care of himself first; what counts is money and power; Striving for personal success is important; Always pursue personal pleasure. Ethnocentrism takes eight 5-point-scale items: Belgium should not have admitted guest workers; Immigrants cannot be trusted; Guest workers threaten the employment of Belgians; Guest workers exploit the social security system; Muslims are a threat to our culture and customs; The presence of different cultures enriches society; Repatriate guest workers when the number of jobs decreases; No political rights and activities for immigrants. Nationalism takes four 5-point-scale-items: Indicate the membership group you feel you belong to (Flemish, Belgian, other); Flanders/Belgium must decide; Belgium has to disappear/strengthened; Split up/federalize social security. Some items of Nationalism was measured in a somewhat more complicated while Authoritarianism is measured by the following two 5-point-scale items: Child has to learn obedience and respect to authority; Solution is to get rid of immoral people. The number of item A used in this study is different from that used in [21]. Because only the two items were used in all of the three times, we only used those two items to measure the latent factor A. Further details on a brief summary of the variables and the sample procedure can be found in [21].

4. RESULTS

As mentioned in introduction, we carried out two stages of modelling, first involving only Individualism (I), Nationalism (N), and Ethnocentrism (E) then adding Authoritarianism (A) to the next stage. The design of all items suggests a simple structure to model the relationship between all items and the latent factors in factor loading matrix. The first four items in first column load on Individualism, the fifth until eight items in second column load on Nationalism, the ninth until sixteenth items in third column load on Ethnocentrism and the remaining two items in last column load on Authoritarianism. The simple structure suggests to set other elements of factor loading matrix will be fixed at 0 thus not allowing any factor rotation. The estimated factor loading ranges from 0.418 to 1.481 for first model and 0.000 to 1.513 for second model, and the variances of unique factors ranges from 0.524 to 10.745 for first model and 0.199 to 9.278 for second model suggest that all latent factors, are well defined by the corresponding items.

For the first model where only three latent factors are in the model (I, N and E), to answer the research question how the latent factors develop over time and whether there is any association between latent factors can be answered by the fixed slope and the correlation matrix of the random effects (see [16]). The slopes for I, N and E are 0.005, 0.440 and -0.141. These results suggest that both I and N increase over time with N increasing faster than I as well as E decrease over time. By adding A in the model, both I (0.005) and N (0.436) still increase over time and N also increasing faster than I. A (0.122) also increases over time but E (-0.109) is consistent with the first model which decreases over time. This shows that by adding A, the development of I, N and E in the model did not change. According to Toharudin et al. [20], by using CT-SEM, I, N and E have a strong tendency to persist over time. By adding A in the CT-SEM model, Angraini et al. [21] also stated that all latent factors have a strong tendency to persist over time. This means that Flanders people will tend to maintain the character of I, N and A and even tend to increase over time based on the results of the LFLMM.

Table 1 Correlation matrix of random effects for I, N, and E

Random effects		I		N		E	
		a_{11}	a_{12}	a_{21}	a_{22}	a_{31}	a_{32}
I	a_{11}	1.000	-0.112	-0.042	0.428	0.598	0.108
	a_{12}	-0.112	1.000	-0.060	0.154	0.011	0.150
N	a_{21}	-0.042	-0.060	1.000	-0.194	-0.108	0.232
	a_{22}	0.428	0.154	-0.194	1.000	0.461	-0.391
E	a_{31}	0.598	0.011	-0.108	0.461	1.000	-0.291
	a_{32}	0.108	0.150	0.232	-0.391	-0.291	1.000

The correlation matrix of random effects for both model are shown in Table 1 and 2. a_{11} and a_{12}, a_{21} and a_{22} , a_{31} and a_{32}, a_{41} and a_{42} are the random intercept and random slope for I, N, E, and A respectively. The correlograms in Figure 1 to 4 visualize the correlation matrices for the random intercept and slope. The coefficient correlation is coloured according to the value. Positive correlation are displayed in blue and negative correlations in red colour. The intensity of the colour is proportional to the coefficient correlation so the stronger the correlation (i.e., the closer to -1 or 1), the darker the circles. The colour legend on the right hand side of the correlogram shows the coefficient correlation and the corresponding colours.

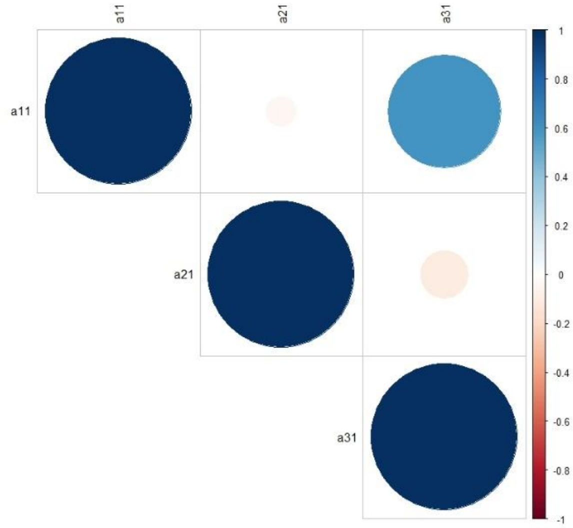

Figure 1 Correlogram of random intercept for I, N, and E

In the first stage model, the positive coefficient correlation between the random intercept of I and E (a_{11} and a_{31}) is 0.598 and shown by the blue circle (Figure 1). It suggests that those who start with better I tend to start with better E. The negative coefficient correlation between the random intercept of I and N (a_{11} and a_{21}) is -0.042 and shown by the small light orange circle (Figure 1). It means that those who start with better I tend to start with worse N. The negative coefficient correlation between the random intercept of N and E (a_{21} and a_{31}) is -0.108. The circle size of the random intercept of N and E is bigger than the random intercept of I and N (Figure 1). The coefficient correlation of the random intercept of N and E suggests that those who start with better N tend to start with worse E.

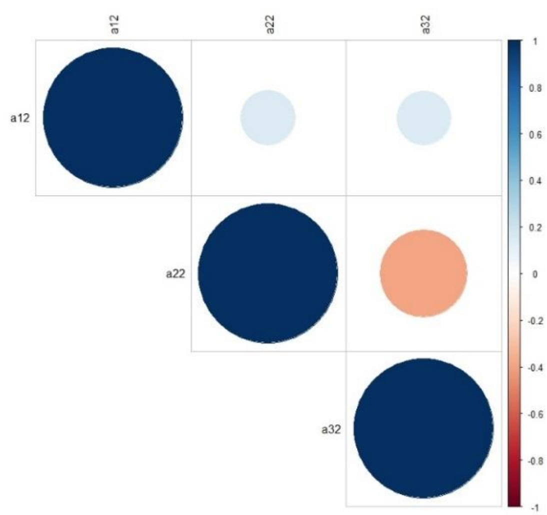

Figure 2 Correlogram of random slope for I, N, and E

The negative coefficient correlation between random slope N and E (a_{22} and a_{32} is shown by the orange circle in Figure 2 and the coefficient correlation is -0.391 (Table 1). It suggests that the development of N and E is related. That is, if one subject's N decreases over time, then it is reasonable to expect that his or her E would also increase over time. The positive coefficient correlation between random slope a_{12} and a_{22}, a_{32}, 0.154 and 0.150, suggest that the development of I and N,E, are related. If one subject's I decreases over time, then it is reasonable to expect that his or her N and E would also decrease over time. These are also shown by the small light blue circle in Figure 2.

Table 2 Correlation matrix of random effects for I, N, E and A

Random effects		I		N		E		A	
		a_{11}	a_{12}	a_{21}	a_{22}	a_{31}	a_{32}	a_{41}	a_{42}
I	a_{11}	1.000	-0.113	-0.042	0.432	0.620	0.097	0.595	-0.083
	a_{12}	-0.113	1.000	-0.058	0.135	-0.008	0.142	-0.057	0.091
N	a_{21}	-0.042	-0.058	1.000	-0.193	-0.074	0.150	0.091	-0.019
	a_{22}	0.432	0.135	-0.193	1.000	0.436	-0.072	0.284	0.130
E	a_{31}	0.620	-0.008	-0.074	0.436	1.000	-0.142	0.672	0.084
	a_{32}	0.097	0.142	0.150	-0.072	-0.142	1.000	0.374	0.369
A	a_{41}	0.595	-0.057	0.091	0.284	0.672	0.374	1.000	-0.159
	a_{42}	-0.083	0.091	-0.019	0.130	0.084	0.369	-0.159	1.000

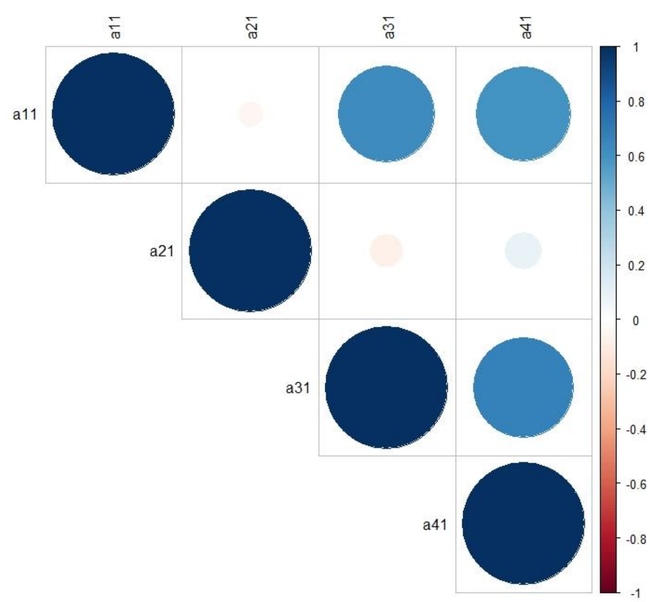

Figure 3 Correlogram of random intercept for I, N, E and A

By adding A in the model, the coefficient correlation between the intercept of I and N (a_{11} and a_{21} , -0.042), I and E (a_{11} and a_{31}, 0.620), N and E (a_{21} and a_{31}, -0.074) in Table 2, also have the same direction with the model without A (Table 1). While the coefficient correlation between random intercept of I and A (a_{11} and a_{41}, 0.595), N and A (a_{21} and a_{41}, 0.091), E and A (a_{31} and a_{41}, 0.672) are also positive and are shown by the blue circle in Figure 3. It suggests that those who start with better I, N and E tend to start with better A. The highest coefficient correlation is E and A. It means there is highest association between E and A. This result is consistent with that obtained in Angraini et al. [21] using the standardized cross-effect in CT-SEM, there are reciprocal effects between E and A although the number of items from A used in this study was only two.

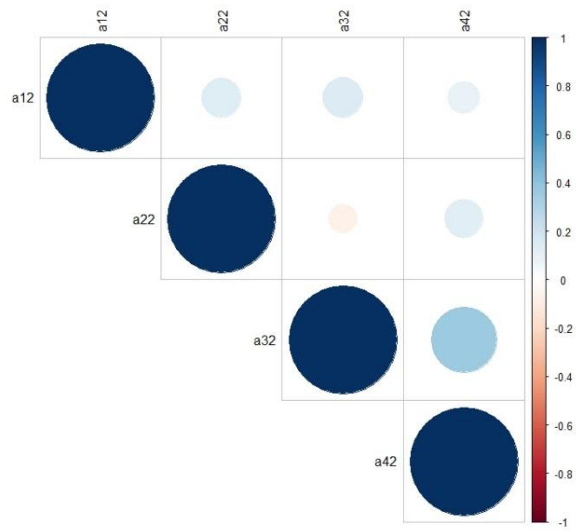

Figure 4 Correlogram of random slope for I, N, E and A

The coefficient correlation between the slope of N and E (a_{22} and a_{32}, -0.072), I and N (a_{12} and a_{22}, 0.135), I and E (a_{12} and a_{32}, 0.142), also have the same direction with the model without A. The coefficient correlation between the slope of I and A (a_{12} and a_{42}, 0.091), N and A (a_{22} and a_{42}, 0.130), E and A (a_{32} and a_{42}, 0.369) are positive (Figure 4). If one subject's I, N, and E decreases over time, then it is reasonable to expect that his or her A would also decrease over time.

5. CONCLUSIONS

In this paper, the LFLMM has been applied to analyse the Flanders data. It has been shown that the LFLMM can accommodate changes over time in high-longitudinal data especially to handle multiple outcomes. It is also shown that I, N and A increase over time while E decreases over time. In line with the results of CT-SEM, Flanders people will tend to maintain the character of I, N, and A in a long period of time.

In summary, the results of the LFLMM analysis to answer whether there is any association between I, N, E and A in Flanders developments demonstrated that there were positive correlation between E and A; I and E; and I and A. This implied that Flanders people who start with better E tend to start with better A. Furthermore, Flanders people who start with better I tend to start with better E and A.

According to An et al. [16], in the LFLMM model, it is possible to add fixed variables to the multivariate linear mixed model for identification of variables affecting the changes patterns over time. For example, in Flanders data, sex and education variables can be taken into account in the model to further understand whether the two variables influence the change patterns in latent factors over time. This idea is recommended for further research using the Flanders data.

ACKNOWLEDGEMENTS

We would like to express our appreciation to the editors and reviewers for their helpful critiques and suggestions. We would also like to thank Jaak Billiet who supported the general election data set.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests.

REFERENCES

[1] G. Verbeke, S. Fieuws, G. Molenberghs, M. Davidian, The analysis of multivariate longitudinal data: A review, Stat. Methods Med. Res. 23 (2014), 42–59.

[2] S. Fieuws, G. Verbeke, Pairwise Fitting of Mixed Models for the Joint Modeling of Multivariate Longitudinal Profiles, Biometrics. 62 (2006), 424–431.

[3] S. Fieuws, G. Verbeke, F. Boen, C. Delecluse, High dimensional multivariate mixed models for binary questionnaire data, J. R. Stat. Soc. C. 55 (2006), 449–460.

[4] S. Fieuws, G. Verbeke, B. Maes, Y. Vanrenterghem, Predicting renal graft failure using multivariate longitudinal profiles, Biostatistics. 9 (2008), 419–431.

[5] S.A. Sivo, Multiple Indicator Stationary Time Series Models, Struct. Equ. Model. 8 (2001), 599–612.

[6] B.O. Muthén, P.J. Curran, General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation., Psychological Methods. 2 (1997), 371–402.

[7] T. Toharudin, Individualism, nationalism, ethnocentrism and authoritarianism: Evidence from Flanders by means of structural equation modeling. Rijksuniversiteit Groningen, 2010.

[8] K.A. Bollen, P.J. Curran, Latent curve models: a structural equation perspective, Wiley-Interscience, Hoboken, N.J, 2006.

[9] H. Singer, Continuous panel models with time dependent parameters, J. Math. Soc. 23 (1998), 77–98.

[10] J.H.L. Oud, R.A.R.G. Jansen, Continuous time state space modeling of panel data by means of sem, Psychometrika. 65 (2000), 199–215.

[11] J. H. L. Oud and M. J. M. H. Delsing, Continuous Time Modeling of Panel Data by means of SEM, Longitudin. Springer, New York, 2010.

[12] M.C. Voelkle, J.H.L. Oud, E. Davidov, P. Schmidt, An SEM approach to continuous time modeling of panel data: Relating authoritarianism and anomia., Psychological Methods. 17 (2012), 176–192.

[13] R. Herliansyah, I. Fitria, Latent variable models for multi-species counts modeling in ecology, Biodiversitas. 19 (2018), 1871–1876.

[14] R.E. Caraka, S. Shohaimi, I.D. Kurniawan, R. Herliansyah, A. Budiarto, S.P. Sari, B. Pardamean, Ecological Show Cave and Wild Cave: Negative Binomial Gllvm's Arthropod Community Modelling, Proc. Comput. Sci. 135 (2018) 377–384.

[15] R. E. Caraka, N. T. Nugroho, S. Tai, R. C. Chen, T. Toharudin, and B. Pardamean, Feature importance of the aortic anatomy on endovascular aneurysm repair (EVAR) using Boruta and Bayesian MCMC, Commun. Math. Biol. Neurosci. 2020 (2020), Article ID 22.

[16] X. An, Q. Yang, and P. M. Bentler, A latent factor linear mixed model for high-dimensional longitudinal data analysis, Stat. Med. 32 (2013), 4229–4239.

[17] M. Noh, Y. Lee, J.H.L. Oud, T. Toharudin, Hierarchical likelihood approach to non-Gaussian factor analysis, J. Stat. Comput. Simul. 89 (2019), 1555–1573.

[18] R. E. Caraka et al., Variational Approximation Multivariate Generalized Linear Latent Variable Model in Diversity Termites, SYLWAN, 164(2020), 161–177.

[19] C. Kondaurova, Maria V;Bergeson, Ronya R; Xu, Huiping; Kitamura, Affective Properties of Mothers' Speech to Infants With Hearing Impairment and Cochlear Implants, J. Speech Language Hear. Res. 58 (2015), 590–601.

[20] T. Toharudin, J. H. L. Oud, and J. B. Billiet, Assessing the relationships between Nationalism, Ethnocentrism, and Individualism in Flanders using Bergstrom's approximate discrete model, Stat. Neerlandica, 62 (2008), 83–103.

[21] Y. Angraini, T. Toharudin, H. Folmer, and J. H. L. Oud, The Relationships between Individualism, Nationalism, Ethnocentrism, and Authoritarianism in Flanders: A Continuous Time-Structural Equation Modeling Approach, Multivar. Behav. Res. 49 (2014), 41–53.

[22] J. Roy and X. Lin, Latent Variable Models for Longitudinal Data with Multiple Continuous Outcomes, Biometrics, 56 (4) (2000), 1047–1054.

[23] A. P. Dempster, N. M. Laird, and D. B. Rubin, Maximum Likelihood from Incomplete Data via the EM Algorithm A., J. R. Stat. Soc. Seri. B, 39 (1977), 1–38.

[24] Interuniversitair steunpunt politieke-opinieonderzoek, General Election Study: codebook and questionnaire, ISPO, Leuven, 1991.

[25] Interuniversitair steunpunt politieke-opinieonderzoek, General Election Study: codebook and questionnaire, ISPO, Leuven, 1995.

[26] Interuniversitair steunpunt politieke-opinieonderzoek, General Election Study: codebook and questionnaire, ISPO, Leuven, 1999.