

2nd International Conference on Computer Science and Computational Intelligence 2017, ICCSCI
2017, 13-14 October 2017, Bali, Indonesia

Linear Mixed Model for Analyzing Longitudinal Data: A Simulation Study of Children Growth Differences

Novi Hidayat Pusponegoro^a, Ro'fah Nur Rachmawati^b,

Khairil Anwar Notodiputro^{c*}, Bagus Sartono^c

^a*Institute of Statistics, Otista Raya No. 64C, Jakarta 13330, Indonesia*

^b*Bina Nusantara University, Kebun Jeruk Raya No. 27, Jakarta 11530, Indonesia*

^c*Bogor Agricultural University, Jl. Raya Darmaga Bogor, Bogor 16680, Indonesia*

Abstract

Growth developmental research is one example of the application of longitudinal data that have correlated value over time. Linear Mixed Model (LMM) is an extension of classic statistical procedures that provides flexibility analysis in correlated longitudinal data and allows researcher to model the covariance structures that represent its random effects. This paper briefly describes growth curves model as a single LMM that represent two levels of observation, which focused on modeling its covariance structure to capture correlated information over time of individual performance. We apply LMM and model different types of its covariance structure in the simulation study of children's growth differences based on the feeding methods. We perform simulation scenario using MIXED procedure in SAS system, based on three fit indices (-2RLL, AIC and SBC) and *p*-value significance level, we obtain Unstructured (UN) covariance is always be the best fit in presenting the characteristic of data but not the best choice considering inefficient numbers of parameters while Heterogeneous First-order Autoregressive (ARH(1)) is a proper alternative covariance structure with ease of data interpretation from fewer numbers of estimated parameters.

© 2017 The Authors. Published by Elsevier B.V.

Peer-review under responsibility of the scientific committee of the 2nd International Conference on Computer Science and Computational Intelligence 2017.

Keywords: covariance structure; repeated measurement; *within-subject*; growth curves; marginal model.

* Corresponding author. Tel.: +62- 251 - 8624535; fax: +62- 251 - 8624535.

E-mail address: khairilnotodiputro@gmail.com

1. Introduction

Linear model is an approach for modeling the relationship between n independent observations $\mathbf{y} = (y_1, \dots, y_n)^T$ with one or more explanatory variables denoted $\mathbf{X} = (x_{ij})$ as $n \times p$ matrix model, where x_{ij} is the value of explanatory variable j for observation i . So that, in matrix notation the linear model can be expressed as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (1)$$

where $\boldsymbol{\beta}$ is a $p \times 1$ parameter vector with $p \leq n$ and $\boldsymbol{\varepsilon}$ is a Gaussian random component having $E(\boldsymbol{\varepsilon}) = \mathbf{0}$ and $\mathbf{V} = \text{Var}(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{I}$. Note that in this classical linear model, the variance structure is presumed in ordinary way that are constant variance and uncorrelated data over time. When we work on longitudinal data analysis this model become less appropriate, because by assuming the constant variance and ignoring the correlation structure the statistical inference from linear model might be erroneous.

Linear Mixed Model (LMM) or Gaussian Mixed Model (MM) analysis is a statistics procedure that provides a flexible approach in statistical analyses with correlated longitudinal data. Longitudinal data is described by response variable with repeatedly measured for each unit analysis and time as the repeated factor. In longitudinal data structure, the number of repeated measure may not be equal for each unit analysis and the time of observation may also not in the same spacing or intervals. In other words, longitudinal data characteristics allow for unequal space and missing value observations. LMM covers characteristics of longitudinal data analysis while in Linear Model are hardly applicable.

Growth developmental researchs evolve in wide variety of scientific fields, such as psychology¹⁰, economics³ and educations¹¹. Growth curves in developmental study are usually represented by means of the two-level hierarchical linear model, in which population parameters, individual effects, and *within-subject* variation are defined at the first level and *between-subject* variation defined at the second level¹³. Growth curves are very interested to develop because we can analyze simultaneously individual level effect/individual differences (or *within-subject*) and group level effect (or *between-subject*)² that changes over time. Harville D. A. (1977)⁷ and Jaume A. et al. (2010)⁸ stated that LMM as the integration of two levels (hierarchical) observation (i.e. *within* and *between-subject*) in a single model, so that LMM fulfill the needs of capturing the individual performance information over time accurately because it allows a variety of variance covariance structure or correlation patterns to be explicitly modeled.

As the specification of Generalized Linear Mixed Model (GLMM), the term Mixed in LMM refers to the fixed and random effect simultaneously used and analyzed in the model. The dissimilarities between fixed and random effect are defined by Lee, Y. et al. (2006)⁹ that fixed effects can describe systematic mean patterns such as treatment methods, while random effects may describe correlation patterns between repeated measures within subjects or heterogeneities between subjects or both. The correlation can be represented by random effects of subjects, and heterogeneities by saturated random effects. In practice, it is often necessary to have both types of random components. However, what really distinguish the LMM from the classic linear model is it models variance and covariance (random effects), so that the parameters of covariance are employed to model data with specific characteristics.

As part of growth developmental research, this paper briefly describes growth curves as a single LMM that represent two-level observations of longitudinal data. Level-1 presents growth shape to capture within-subject effect and level-2 presents growth parameters that characterized between-subject differences. As we stated before, the excellence of LMM is the variance covariance modeling while in linear model it is presumed as uncorrelated and constant variance over time. In this paper, we model the variance and covariance using different structure types such as Unstructured (UN), Compound Symmetric (CS), Heterogeneous Compound Symmetric (CSH), First-order Autoregressive (AR(1)) and Heterogeneous First-order Autoregressive (ARH(1)). In orders to explore the characteristic of correlated longitudinal data and excavate the individual trend performance, we perform simulation study and apply variance covariance modeling to generated children's weight growth that we divide into two groups based on its feeding methods. Simulation study developed using MIXED procedures in SAS system that allows researcher to analyze both fixed and random effect in LMM.

2. Two-Level LMM in Growth Curves Analysis

LMM also known as multi-level modeling^{5,6} and for longitudinal data LMM is the integration of *Within-group* or *Within-subject* model and *Between-group* or *Between-subject* model. At the first level, linear regression of the observations is defined and the growth patterns of individual are modeled as a function of a single set of parameters. Usually, the response variable is modeled as a function of time and a *within-subject* residual component. In matrix notation, the level-1 *Within-subject* model is:

$$\mathbf{y}_i = \mathbf{T}_i \boldsymbol{\beta}_i + \mathbf{e}_i, \mathbf{e}_i \cong N(0, \mathbf{R}). \quad (2)$$

Within-subject model is defined that y_i is function of β_i and T_i where y_i is $t \times 1$ repeated-measure vector of the subject i , and β_i is a $p \times 1$ individual vector parameter which specifies the shape of the growth curve for subject i , and T_i is the $t \times p$ matrix of time and their polynomial transformation. It is assumed that the errors are independent with normal multivariate distribution $N(0, \mathbf{R})$. Thus, $\mathbf{R} = \text{Var}(\mathbf{e}_i)$ is a positive-definite covariance matrix. However, when the observations have a given order or a specific structure, it should be assumed that correlation between residuals is distinct from zero and varies systematically.

At the second level, growth parameters are modeled as a function of population averages and the deviations of subjects from these parameters. In the longitudinal context, the level-2 *between-subject* model takes the individual growth parameters (β_i) as random dependent variables, using matrix notation the model can be expressed as follows:

$$\boldsymbol{\beta}_i = \mathbf{Z}_i \boldsymbol{\gamma} + \mathbf{u}_i, \mathbf{u}_i \cong N(0, \mathbf{G}). \quad (3)$$

Formula above explain that the subject growth is functions of Z_i is the $p \times q$ *between-subject* design matrix with known and fixed items $\boldsymbol{\gamma}$ is a $q \times 1$ fixed-coefficient vector and \mathbf{u}_i is a $p \times 1$ random-error vector with zero mean and variance \mathbf{G} .

The term of fixed and random effects in LMM for longitudinal data defined that fixed effects have levels that are of primary interest and would be used again if the experiment were repeated. Random effects have levels that are not of primary interest, but rather are thought of as a random selection from a much larger set of levels. Subject effects are almost always random effects, while treatment levels are almost always fixed effects. Other examples of random effects include cities in a multi-site trial, batches in a chemical or industrial experiment, and classrooms in an educational setting. By substituting formula (3) in formula (2) we had formula (4) as matrix notation of LMM:

$$\mathbf{Y}_i = \mathbf{TZ}_i \boldsymbol{\gamma} + \mathbf{Tu}_i + \mathbf{e}_i, \quad (4)$$

where \mathbf{Y}_i is the $t \times 1$ repeated-measure vector of the subject i , and $\mathbf{Z}_i \boldsymbol{\gamma} + \mathbf{u}_i$ is a term that specifies the shape of the growth curve for subject i , called growth parameter β . Furthermore, the $\mathbf{TZ}_i \boldsymbol{\gamma}$ term is fixed effect and the $\mathbf{Tu}_i + \mathbf{e}_i$ is random effect. It is supposed that the level-1 error $\mathbf{e}_i \cong (0, \sigma^2 \mathbf{I})$ and the level-2 error $\mathbf{u}_i \cong N(0, \mathbf{G})$. The between-subject model complexity increases when covariate which is needed to explain variation between growth parameters is included.

3. Modeling Variance Covariance Structures

3.1. The Structure of Variance Covariance Matrix

Some standard statistical techniques, such as the unpaired *t-test*, simple linear regression or the *chi-square test* for association, assume that each of the primary observations that presume a data set as independent among others^{1,4}. Unfortunately, this assumption can be inappropriate if repeated longitudinal data observations are taken within subjects because observations within an individual tend to be correlated over time with one another. It is assumed that the probability distribution of repeated measurements has the same shape for each subject, but the distribution parameters vary among subjects. This means that each repeated observation in an individual may provide less additional information than a new observation in a new individual. It is absolutely known, that if a standard statistical analysis which assumes all observations to be independent is performed on repeated measures data when the intra-class correlation is positive, so the results may be misleading.

One of the LMM advantages is modeling variance covariance matrix structure of data denote as \mathbf{R} , since the matrix determines pattern of autocorrelation among residual components. The modeling of \mathbf{R} structure gives estimation improvement accuracy for the regression parameter and test power for detection treatment effects.

If Gaussian data observations are measured over time on the same subject then *within-subject* correlation naturally becomes part of the **R** matrix or the conditional variance of the observations given the random effects. However, the observations at every time have a unique variance and observations within the same subject at every pair of times have a unique covariance. For J times observation, there are $J(J+1)/2$ variance covariance parameters forming σ in the Unstructured (UN) matrix and it is not an efficient model considering the numbers of estimated parameters. UN is the most typical **R** structure for longitudinal data, since it requires no assumption regarding the error terms and allows unique correlation pattern between observations.

Table 1. Covariance Structures for J times observation

Type	Matrix R	Number of Parameters
Unstructured (UN)	$UN = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1J} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{J1} & \sigma_{J2} & \sigma_{J3} & \sigma_J^2 \end{bmatrix}$	$J(J+1)/2$
Compound Symmetric (CS)	$CS = \sigma^2 \begin{bmatrix} 1 & \rho & \cdots & \rho \\ & 1 & \cdots & \vdots \\ & & \ddots & \rho \\ & & & 1 \end{bmatrix}$	2
Heterogenous Compound Symmetric (CSH)	$CSH = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \cdots & \sigma_1\sigma_J\rho \\ & \sigma_2^2 & \cdots & \sigma_2\sigma_J\rho \\ & & \ddots & \vdots \\ & & & \sigma_J^2 \end{bmatrix}$	$J+1$
First-order Autoregressive (AR(1))	$AR(1) = \sigma^2 \begin{bmatrix} 1 & \rho & \cdots & \rho^{J-1} \\ & 1 & \cdots & \rho^{J-2} \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix}$	2
Heterogenous First-order Autoregressive (ARH(1))	$ARH(1) = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \cdots & \sigma_1\sigma_J\rho^{J-1} \\ & \sigma_2^2 & \cdots & \sigma_2\sigma_J\rho^{J-2} \\ & & \ddots & \vdots \\ & & & \sigma_J^2 \end{bmatrix}$	$J+1$

We can explore the types of **R** structures for J times observation in Table 1. Note that each diagonal term presents the variance of level-1 repeated observations each time, elements in above diagonal term present the covariance of level-1 repeated observation between times and of course all matrices are square symmetric. AR(1) and CS assume that all variances over time are constant, with constant correlation over time in CS while exponentially decrease correlation in AR(1). AR (1) defines explicitly higher correlation in adjacent time and smaller correlation for larger time intervals. Another possible covariance model as generalizations of CS and AR(1), that is, CSH and ARH (1) respectively, which have unique value of variance each time. Since AR(1), CS, ARH(1) and CSH structure model have the constant parameter, then these can be more efficient and powerful **R** structures in terms of detecting treatment effects.

3.2. Testing and Fitting **R** Structure

Testing of the matrix form **R** indispensable, this ensures that covariance matrix modeling is appropriate. The likelihood ratio test for covariance components is similar to the likelihood ratio test for estimable functions, we define the null hypothesis for **R** is ordinary form, i.e. $H_0: \sigma = \sigma_0$, and test this null hypothesis compared to its full

model. Since, the model with Gaussian data as an LMM then estimation of σ using Restricted Maximum Likelihood (REML) method so that the resulting of likelihood ratio statistic can be written as follows:

$$LR(\sigma_0) = -2 \log \left[\frac{f_R(\hat{\sigma}_0)}{f_R(\hat{\sigma})} \right], \quad (5)$$

where $f_R(\hat{\sigma})$ denotes the REML likelihood.

The most important application of **R** testing is its use for selecting among competing covariance models. Fitting **R** structure model can be conducted graphically that represent the information of data or by comparing the **R** estimates using fit indices criteria which can provide alternatives to formal testing¹². In the latter option the most widely used criteria to select the best **R** structure with models that have the same mean structure are Negative 2-Residual Log-likelihood (-2RLL), Akaike's Information Criterion (AIC) and Schwarz's Bayesian Criterion (SBC). -2RLL = -2 Log (L); compares the residual produce from full model (when all parameters estimated) and interested/reduced model. The AIC = -2 Log (L) + 2 d ; statistics can be interpreted as a maximization of the expected entropy, and BIC = -2 Log (L) + p Log (N); statistics generally penalizes free parameters more strongly than AIC. Note that L states the log-likelihood function, p denotes the number of model parameters that are estimated and N is the sample size.

4. Simulation Study: R Structure Models for Children's Growth Differences

In this paper, **R** modeling structure focus to determine the correlation patterns between repeated measurements using the LMM approach of longitudinal data especially in its marginal forms. Marginal terms correspond to level-1 *within-subject* random effect, while level-2 random effect is not a concern of this research. The correlation patterns can be emphasized form the fitted **R** structure among the linear model, unique variance covariance value model, constants variance over time model or heterogeneous structured model.

We perform simulation study using MIXED procedure in SAS system to model **R** structures. SAS have several procedures which offer programming capabilities to analyze mixed model including NLMIXED, GLIMMIX and MIXED procedures. Based on correlated longitudinal data and variety of **R** structures modeling, MIXED fits the **R** structure to the Gaussian data by using REML as default method estimation. The MIXED procedure provides model flexibility not only the means of data (as in the standard linear model) but also their variances-covariance as well.

4.1. Generated Children's Growth Data

We generate children's weight as our longitudinal data to examine whether there were systematic inter-individual differences in intra-individual change in children's weight over time as a consequence of the differences in feeding methods which are breast and bottle feeding. The breast-fed group comprising 56 and the bottle-fed group consist of 84 participants. Weight (in kilos) is used as dependent variable was measured at birth and on five subsequent occasions i.e. at 1, 2, 3, 4, and 5 years. So, we have, the same number of observation (i.e. five times weight observation for each child) and the same interval of observation (i.e. one-year interval for each child) in our longitudinal data context.

The children's weight generated based on developmental study in Jaume A. et al. (2010)⁸ by the mean weight values over time according to the types of feeding. The observation informs that mean weight values over 5 years respectively are 10, 13.5, 15, 17, 18.5 for the bottle-fed group and 9.8, 12.8, 14.2, 16, 17.4 for the breast-fed group. From the mean weight, we know that bottle-fed dominate the weight in each year observation, although breast-fed always have smaller mean weight than the bottle-fed, these groups have same increasing pattern over time. Generated children's weight consisted of 140 observation per cohort with assumption that random errors e_i are independent with normal multivariate distribution $N(0, \mathbf{R})$ thus children's weight data are also multivariate normal distribution with mean matrix $\boldsymbol{\mu} = \{10 \ 13.5 \ 15 \ 17 \ 18.5; 9.8 \ 12.8 \ 14.2 \ 16 \ 17.4\}^t$ and have Unstructured **R** matrix. Description the within-subject profile based on feeding method are shown in figure 1, its show that weight increasing during the first year of life and moreover children's weight follows increasing linear trend for both groups. The pattern in figure 1 also indicates weight variances are getting higher over time. Both groups show larger variance in weight along with the children's age, larger variance occurs in the fifth years of observation.

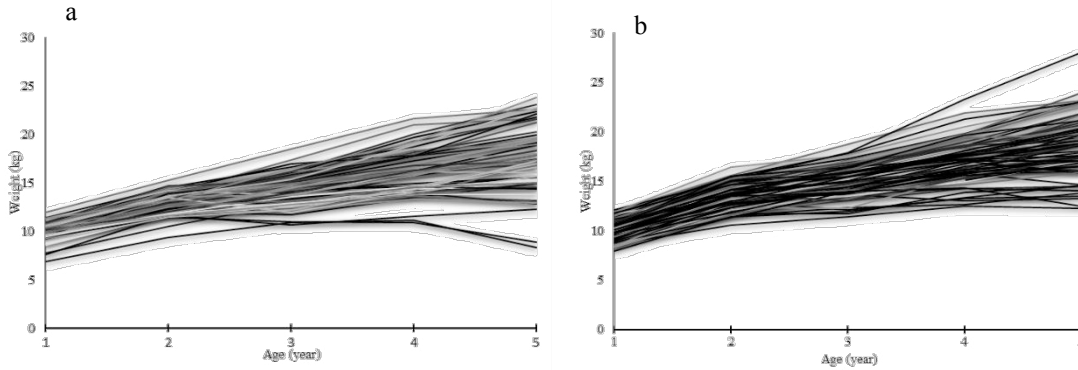


Fig. 1. (a) Breast-Fed Group Profile; (b) Bottle-Fed Group Profile.

4.2. R Structures for Generated Children's Growth Data

Describing data graphically could be one of the essential things before we model the **R** structures. As we can see in Figure 1, when our data indicates children's weight variances increase in every year thus modeling **R** structures with constant variances model such as AR(1) and CS are no longer relevant. Using these constant **R** structures can eliminate the information of children's weight differences each time and its correlation over time, so that non-constant variance model such as UN, ARH(1) and CSH(1) could be the alternatives to capture the variability of within-subject weight over time according to children's weight characteristic described in figure 1.

When modeling developmental change in children's growth, it is important to use model that can capture the increase variability and correlation over time. As mentioned above, MIXED procedure in SAS system fulfill the requirement to model the covariance structure, with the following command:

```
proc mixed data=weight covtest ic;
class feed age children;
model weight = feed age feed*age/s noint;
repeated / type=arh(1) subject=children r rcorr;
run;
```

The ARH(1) matrix defines in TYPE=ARH(1) using REPEATED statement, CLASS defines all variables that interested in the model, individual effect or in this case are the level-1 observation defined in SUBJECT=CHILDREN, and the above command display the ARH(1) matrix and its correlation each time with command R and RCORR respectively. The UN and CSH matrix also can be specified with MIXED procedure using different command TYPE in its REPEATED statement and summarized in Table 2.

Table 2. Estimated **R** Structure for CSH and ARH(1).

R Structure	Age				
	1	2	3	4	5
CSH	1	0.7156	0.7156	0.7156	0.7156
	0.7975	1.2662	2.1570	4.0891	8.4797
ARH(1)	1	0.8545	0.7301	0.6239	0.5331
	1.3136	1.7949	2.4844	3.5536	8.6884

Note in Table 2, correlation of children's weight in the first line and its variance in the second line. Using CSH and ARH(1), we have 5 unique variance each time and 1 correlation parameter, in CSH $\rho = 0.7156$ present constant correlation over time, while in ARH(1) $\rho = 0.8545$ present exponentially decreasing correlation over time. Although not stated in this paper, UN structure requires 15 unique parameters to describe children's growth characteristic. All structures of **R** successfully capture the increase variability of children's weight in five times observation.

4.3. Best Fitted \mathbf{R} Structures

Testing the ordinary structure of \mathbf{R} , i.e. $\mathbf{R} = \sigma^2 \mathbf{I}$, is one of the convincing procedures in longitudinal data that have non-constant variance and correlated over time, the result of this hypothesis stated in table 3. The null hypothesis is statistically significant ($p < .0001$) for all \mathbf{R} structures, in other words we can say that the variances of children's weight are not constant and have high correlation over time, so that all \mathbf{R} structure have better fit rather than \mathbf{R} in ordinary form. However, which model gives the better fit for \mathbf{R} structure of children's growth differences? We can evaluate this using three criteria produced from PROC MIXED in SAS system, these are: -2RLL, AIC and SBC.

In orders to obtain best fitted \mathbf{R} structure and to generalize result of this simulation study, we explore deeper simulation scenarios, i.e. we generate children's weight data with 100 replications build from UN, CSH and ARH(1). In each repetition, we model the \mathbf{R} structure using different type non-constant variance models, for example if we build with UN structure, then we model the data using UN its self and two different types i.e. CSH and ARH(1). From 100 replications, we concern in the average value of three fit indices criteria: -2RLL, AIC and SBC, based on these criteria we rank the model which produce smaller value, the smaller the value the smaller the rank. We sum the rank produce in each \mathbf{R} structure model, so that, the smaller the rank yield the better fit for \mathbf{R} structure. This simulation scenario also summarized in table 3. From 100 replications of children's weight data, UN structure has the smaller rank in three criteria at once. This is not the surprising result, because UN with 15 parameters describes children's weight data properly, however for simplicity interpretation and efficient numbers of parameters, UN does not fulfill these requirements so we purpose the alternative structures of \mathbf{R} . From table 3, ARH(1) is in the second place after UN, so ARH(1) could be one of its alternative structures.

Table 3. Rank of Fit Statistics and Null Model Likelihood Ratio Test.

Criteria	\mathbf{R} Structure Model		
	CSH	ARH(1)	UN
-2RLL	8	7	3
AIC	7	6	5
SBC	7	6	5
<i>p-value</i>	<0.0001	<0.0001	<0.0001

Table 4. *P-values* of feeding effect for different \mathbf{R} structures.

Generated \mathbf{R} Structure	\mathbf{R} Structure Model		
	ARH(1)	CSH	UN
ARH(1)	0.0581	0.0591	0.0585
CSH	0.0358	0.0417	0.0418
UN	0.0508	0.0539	0.0544

PROC MIXED also produces test of fixed effect, this test result *p-value* that represent the significant meaning of the effect or treatment. We use *p-value* as additional criteria to decide which \mathbf{R} structure that gives the best fitted to present children's weight characteristic over time. The smaller the *p-value*, the more significant the effect and gives higher accuracy for estimate regression parameters. In children's weight data we have FEED, AGE and the interaction FEED*AGE as the fixed effects, since feeding method is the concerned covariate in children's growth weight differences, therefore it is important to examine the effect. Table 4 summarizes *p-value* results in PROC MIXED procedure, and we obtain that for any generated \mathbf{R} structure, ARH(1) model gives smallest *p-value* for fixed FEED effect rather than other models. In other words, based on three fit indices criteria (-2RLL, AIC and SBC) and *p-value* significance level, ARH(1) is the best alternative in modeling \mathbf{R} structure to describe children's growth differences based on differences in feeding methods.

5. Conclusions

This paper describes LMM as a single growth curve model that represent two levels of observation. Level 1 describes individual growth shapes as a function of time and its random error which assumed has normal distribution $N(0, \mathbf{R})$, while Level 2 represents growth parameters as a function of fixed effects and its random error also assumed has normal distribution $N(0, \mathbf{G})$. This research focus in modeling level 1 random effect using different types of variance covariance structures. It is essential to describe data graphically before modeling the \mathbf{R} structure, in graphical display it is easier to capture the information whether our longitudinal data has constant

variance over time or vice versa. Based on this graphical information, we can decide to model **R** structure using constant variance over time such as CS and AR(1), or using non-constant variance such as UN, CSH and ARH(1).

We explore how differences in feeding effects the children's weight using LMM. As our longitudinal data, we generate 140 children's weight (in kilos), divided into two groups that characterized the differences in feeding methods i.e. breast-fed and bottle-fed. From graphical information which shows non-constant variance over time, we model **R** structure using UN, CSH and ARH(1) in SAS system with MIXED procedure. The simulated scenarios obtain UN with 15 parameters perfectly describes children's growth correlation and variability over time. When fitting data with statistical models, there is often a compromise between model adequate fit and simplicity. Although UN has the best fit, it gives inefficient numbers of estimated parameters which hardly interpret when we have large times of observation. Under the need of ease interpretation, efficient numbers of estimated parameters, based on three fit indices criteria (-2RLL, AIC and SBC) and *p*-value of feeding, we purpose ARH(1) is a proper **R** structure to describe children's growth differences.

There are many types of **R** structure can be described in longitudinal data analysis, such as Huynh-Feldt, Autoregressive Moving Average, and Ante-dependence. Exploring other different types of **R** structure could be an interested topic for further discussion. This paper concerns LMM in marginal form, in other words we model **R** structure as the part of the level 1's random effect. Although not the scope of this work, the complexity of LMM increase when **G** structure as level 2's random effect can also be modeled.

References

- [1] Armitage, P., and Berry, G. (1987) "Statistical methods in medical research, 2nd ed.", Oxford: Blackwell Scientific Publications.
- [2] Daniel M, James A. D., and James D. M. (2008) "Statistical and computational models of the visual world paradigm: Growth curves and individual differences." *Journal of Memory and Language*, **59**: 475-494.
- [3] Elisabetta A., and Francesco B. (2017) "Children's multidimensional health and medium-term cognitive skills in low- and middle-income countries" *The European Journal of Development Research*, **29** (2): 289-311.
- [4] Feldman, H. A. (1988) "Families of lines: random effects in linear regression analysis." *Journal of Applied Physiology*, **64**: 1721-1732.
- [5] Goldstein, H. (1987) "Multilevel models in educational and social research." London: Charles Griffin & Company Ltd.
- [6] Goldstein, H. (1988) "Multilevel mixed linear modelling analysis using iterative generalized least squares." *Biometrika*, **73**: 43-56.
- [7] Harville, D. A. (1977) "Maximum likelihood approaches to variance component estimation and to related problems" *Journal of the American Statistical Association*, **72**: 320-340.
- [8] Jaume A., Nekane B., Roser B., and Arantaka G. (2010) "General linear mixed model for analysing longitudinal data in developmental research." *Perceptual and Motor Skills*, **110** (2): 547-566.
- [9] Lee, Y, Nelder, J.A. and Pawitan, Y. (2006) "Generalized linear models with random effects.", Boca Raton: CRC Press.
- [10] Robert, S., and Kofi M. (2014) "Some growth points in African child development research" *New Direction for Child and Adolescent Development Wiley online library*, **146**: 97–112.
- [11] Sarah C and Paul H. (2016) "Representation of academics from developing countries as authors and editorial board members in scientific journals: does this matter to the field of development studies?" *The European Journal of Development Research*, **29** (2): 369-383.
- [12] Stroup, W. W. (2013). "Generalized linear mixed models, modern concepts, methods and applications." Boca Raton: CRC Press.
- [13] Timm, N. H. and Mieczkowski, T. A. (1997) "Univariate and multivariate general linear models: theory and applications using SAS software.", NC: SAS Institute.