

Examination of Covariance Structures for Experimental Design with Repeated Measure

Abraham Okolo¹, Saidu Sauta Abdulkadir¹, Ikeme John Dike¹, Abubakar Adamu^{2,*}

¹Department of Statistics and Operations Research, School of Physical Sciences, Modibbo Adama University of Technology, Yola, Nigeria

²Department of Mathematics, Faculty of Science, Gombe State University, Gombe, Nigeria

Email address:

abumalala22@yahoo.com (A. Adamu)

*Corresponding author

To cite this article:

Abraham Okolo, Saidu Sauta Abdulkadir, Ikeme John Dike, Abubakar Adamu. Examination of Covariance Structures for Experimental Design with Repeated Measure. *American Journal of Theoretical and Applied Statistics*. Vol. 9, No. 3, 2020, pp. 63-73.

doi: 10.11648/j.ajtas.20200903.16

Received: April 3, 2020; **Accepted:** May 3, 2020; **Published:** May 27, 2020

Abstract: This research examine different covariance structures for experimental design with repeated measure data. Multiple responses taken sequentially from same experimental unit at different periods of time for quantitative data are referred as Repeated Measurement. Weight of 105 broilers in grams for six group obtained from jewel farm Gombe were used as research materials/data. Eleven different covariance structures including the modified one (UN, UNC, TOEP, TOEPH, ANTE(1), AR(1), ARH(1), CS, CSH, HF and ARFA(1)) were examined. AIC, AICC, BIC, HQIC, CAIC and the modified criteria ASIC were used to examine covariance structures and bring the best among them using the named information criteria. The result shows that sphericity assumptions was violated a such the best covariance structure was ARH(1) while the least structure was CSH. Also on the basis of goodness of fit criteria HQIC was found to be the best information criteria. When examined the best information criteria and covariance structure with the modified ones, the modified ASIC and ARFA(1) found to be the best. In conclusion examine different covariance structures with repeated measure data give a very good result depending on the kind of data.

Keywords: Covariance Structures. Experimental Design, Repeated Measures, Information Criteria, Sphericity Test

1. Introduction

Repeated Measures Designs (RMD) are popular because they allow subject to serve as their own control [1]. This improves the precision of the experiment by reducing the size of the error variance on many of the F – tests. Repeated measure design has several advantages that make it become popular and active in research. Most important it reduces the unsystematic variability in the data and provide greater power to detect the effects, also it assumed that data from different experimental areas will be related, we assume that the relationship between pairs of experimental conditions is similar [2]. Repeated measures can occur in any common experimental design like completely randomized design, randomized complete block design, split block design, strip plot design and so on with a straightforward analysis of variance procedures to analyze the data if the following

assumptions are valid: *Normality*: the data are normally distributed with mean 0 and the variance σ^2 . *Homogeneity of variance*: the error variances are equal. *Sphericity*: the variances between all pairs of the repeated measurements are equal, the so-called compound symmetry pattern (i.e. the covariance between observations within any two different factor levels be the same). Compound symmetry is a special case of more general property of sphericity. If compound symmetry exists, then sphericity also exists, but it is possible for sphericity to exist when compound symmetry does not. At times, the structure of covariance matrix can be extremely complicated. Second, the data may not be complete due to one reason or another, an incomplete or unbalanced repeated measures design is much more cumbersome to deal with than a design which does not have missing observations [2].

Analysis of covariance structures is the common term for a number of techniques for analyzing multivariate data in order to detect and assess latent sources of variation and

covariation in the observed measurements. The techniques of covariance structure analysis are general and flexible in that they can handle many types of covariance structures useful especially in the behavioral and social sciences. Although these techniques can be used for exploratory analysis, they have been most successfully applied to confirmatory analysis where the type of covariance structure is specified in advance. A covariance structure of a specified kind may arise because of a specified substantive theory or hypothesis, a given classificatory design for the measures, known experimental conditions, or because of results from previous studies based on extensive data. Sometimes the observed variables are ordered through time, as in longitudinal studies, or according to linear or circular patterns, as in [3] simplex and complex models, or according to a given causal scheme, as in path analysis. Several investigators have considered covariance structure analysis as an approach to study differences in test performance when the tests have been constructed by assigning items or subtests according to objective features of content or format to subclasses of a factorial or hierarchical classification [4, 5, 6]. Once you have the random effects determined, then you can move on to selecting the covariance structure. There are a variety of considerations when selecting the covariance structure. They include the number of parameters, the interpretation of the structure, diagnostic results, and effects on fixed effects. If the data suffices, one could always fit the unstructured covariance structure and go with it. However, just as in traditional regression we want to have as few parameters in the model as possible. The more data you have the more parameters you can fit, but they do not always add to our knowledge and often take away. The more complex the model the more specific to the data it will be and the less generalizable. Our belief that nature follows simple, elegant rules leads us to look for the simpler model when possible. From the fixed effects perspective, selecting a structure that is too simple increases the fixed effects, Type I error rate, and selecting a structure that is too complex sacrifices power and efficiency. Covariance measures how much variation in one variable is explained by another variable and is used to calculate correlations. Covariance *structures* describe mathematical patterns exhibited by covariance and correlation matrices. Some covariance structures require that the measurements occur at equally spaced intervals, while others are more flexible and do not need this requirement. Model diagnostics, such as AICC, AIC, BIC and other measures are used to select the covariance structure that best fits the data. Selecting the right covariance structure is not an end unto itself. It is an intermediate step in obtaining correct tests and inference about the fixed effect means. The usual strategy for selecting the best covariance structure, without *a priori* knowledge of a repeated measures process, is to fit the structures appropriate for the measurement points, (equal or unequal intervals) and compare the AICC, AIC or BIC values. (Remember, the information criterion values

are relative and indicate which options provide the better fit. Not how well the model fits the data in absolute terms. This research also examined the impact of using different covariance structures for experimental design with repeated measure data.

2. Materials and Method

2.1. Source of Data

A secondary data was used which was obtained from Gombe farm. The sample size of the data used consist of 105 broilers. The sample was divided into six groups, where each was fed with different kind of food. The weight of each broiler was measured in grams on weekly basis during the period, starting from week 0 to week 6.

2.2. Mauchly's Sphericity Test

[7]Highlighted that when sphericity tests such as the techniques outlined by [8] shows variance inequalities in the sample; the effect of violating sphericity assumption is a loss of power and a test statistic that cannot be compared to tabulated values of the F – distributions. The authors recognized that even if sample variances are unequal, such inequalities might simply reflect sampling error. Therefore, Mauchly's sphericity test (i.e., Mauchly's W) be used to test the null hypothesis that the homogeneity condition holds in the population.

When the sphericity test is significant, SPSS, R or Stata Packages offers two ways to test the significance of the within-subject effects. The first way is to adjust the univariate tests themselves, SPSS, R or Stata packages gives three such adjustments: Greenhouse – Geisser Epsilon adjusted F test which was developed by [9], the less conservative Huynh – Feldt Epsilon adjusted F test which was developed by [10] and the Lower bound Epsilon. The second way involves four different multivariate tests: Wilks' Lambda, Pillai's Trace, Hotelling-Lawley Trace and Roy's Greatest Root. None of these methods has been indicated to be better than the others. Hence, all are similar to using the Hotelling T^2 statistic.

$$Hotelling's T^2 = N (GM)' S_{wg}^{-1} (GM) \quad (1)$$

whereas,

N indicates the number of groups.

GM indicates the overall mean of groups.

Swg indicates the treatments variance-covariance matrix.

Wilks' λ can be obtained using the equation below:

$$\lambda = \frac{1}{1 + T^2} \quad (2)$$

Hypothesis,

H0: The groups mean are significantly equal.

H1: The groups mean are significantly not equal.

Decision rule: Reject Ho if $P < 0.05$ and conclude that there

is significant differences between the groups otherwise accept H_0 if $P \geq 0.05$ and conclude vice versa.

2.3. Covariance Structures

This research shall examine the consequences of fitting models that assumed homogeneous variances to data that exhibit variance heterogeneity. A brief summary of the covariance structures with seven (7) repeated measures are as follows.

2.3.1. Compound Symmetry (CS)

The CS structure is a covariance pattern where each main diagonal element is decomposed to unity and all off-diagonal elements are set to the value of ρ .

$$CS = \sigma^2 \begin{bmatrix} 1 & \rho & \rho & \rho & \rho & \rho & \rho \\ & 1 & \rho & \rho & \rho & \rho & \rho \\ & & 1 & \rho & \rho & \rho & \rho \\ & & & 1 & \rho & \rho & \rho \\ & & & & 1 & \rho & \rho \\ & & & & & 1 & \rho \\ & & & & & & 1 \end{bmatrix} \quad (3)$$

2.3.2. Heterogeneous Compound Symmetry (CSH)

This pattern is a simple generalization of the CS structure that allows for non-constant variances on the main diagonal. While the assumption of sphericity is met in the CS structure, sphericity is not met in CSH due to the introduction of variance heterogeneity.

$$\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho^2 & \sigma_1\sigma_4\rho^3 & \sigma_1\sigma_5\rho^4 & \sigma_1\sigma_6\rho^5 & \sigma_1\sigma_7\rho^6 \\ & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho^2 & \sigma_2\sigma_5\rho^3 & \sigma_2\sigma_6\rho^4 & \sigma_2\sigma_7\rho^5 \\ & & \sigma_3^2 & \sigma_3\sigma_4\rho & \sigma_3\sigma_5\rho^2 & \sigma_3\sigma_6\rho^3 & \sigma_3\sigma_7\rho^4 \\ & & & \sigma_4^2 & \sigma_4\sigma_5\rho & \sigma_4\sigma_6\rho^2 & \sigma_4\sigma_7\rho^3 \\ & & & & \sigma_5^2 & \sigma_5\sigma_6\rho & \sigma_5\sigma_7\rho^2 \\ & & & & & \sigma_6^2 & \sigma_6\sigma_7\rho \\ & & & & & & \sigma_7^2 \end{bmatrix} \quad (6)$$

2.3.5. Toeplitz (TOEP)

The TOEP structure is akin to the AR(1) structure in that the covariance between any two measurement occasions is constant for all at the same time lag.

$$TOEP = \begin{bmatrix} \sigma^2 & \sigma_1 & \sigma_2 & \sigma_3 & \sigma_4 & \sigma_5 & \sigma_6 \\ & \sigma^2 & \sigma_1 & \sigma_2 & \sigma_3 & \sigma_4 & \sigma_5 \\ & & \sigma^2 & \sigma_1 & \sigma_2 & \sigma_3 & \sigma_4 \\ & & & \sigma^2 & \sigma_1 & \sigma_2 & \sigma_3 \\ & & & & \sigma^2 & \sigma_1 & \sigma_2 \\ & & & & & \sigma^2 & \sigma_1 \\ & & & & & & \sigma^2 \end{bmatrix} \quad (7)$$

2.3.6. Heterogeneous Toeplitz (TOEPH)

Once again, this pattern is a generalization that allows for non-constant variances with respect to measurement occasions.

2.3.3. First Order Autoregressive (AR(1))

The AR(1) structure was adapted from time series analysis where it is often used to account for the internal structure or autocorrelation of observations over time. It models the relationships of the measurement occasions parsimoniously by imposing a decreasing exponential trend between time lags.

$$AR(1) = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 & \rho^4 & \rho^5 & \rho^6 \\ & 1 & \rho & \rho^2 & \rho^3 & \rho^4 & \rho^5 \\ & & 1 & \rho & \rho^2 & \rho^3 & \rho^4 \\ & & & 1 & \rho & \rho^2 & \rho^3 \\ & & & & 1 & \rho & \rho^2 \\ & & & & & 1 & \rho \\ & & & & & & 1 \end{bmatrix} \quad (5)$$

2.3.4. Heterogeneous First Order Autoregressive (ARH(1))

This is a simple generalization of the AR(1) pattern that allows for non-constant variances with respect to measurement occasions. This is a first-order autoregressive structure with heterogeneous variances.

$$\text{TOEPH} = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_1 & \sigma_1\sigma_3\rho_2 & \sigma_1\sigma_4\rho_3 & \sigma_1\sigma_5\rho_4 & \sigma_1\sigma_6\rho_5 & \sigma_1\sigma_7\rho_6 \\ & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho_2 & \sigma_2\sigma_5\rho_3 & \sigma_2\sigma_6\rho_4 & \sigma_2\sigma_7\rho_5 \\ & & \sigma_3^2 & \sigma_3\sigma_4\rho & \sigma_3\sigma_5\rho_2 & \sigma_3\sigma_6\rho_3 & \sigma_3\sigma_7\rho_4 \\ & & & \sigma_4^2 & \sigma_4\sigma_5\rho & \sigma_4\sigma_6\rho_2 & \sigma_4\sigma_7\rho_3 \\ & & & & \sigma_5^2 & \sigma_5\sigma_6\rho & \sigma_5\sigma_7\rho_2 \\ & & & & & \sigma_6^2 & \sigma_6\sigma_7\rho \\ & & & & & & \sigma_7^2 \end{bmatrix} \quad (8)$$

2.3.7. Unstructured (UN)

The UN pattern imposes no constraints on the form of the covariance matrix and, as a result, requires a parameter estimate for each element of the matrix.

$$\text{UN} = \begin{bmatrix} \sigma_1^2 & \sigma_{21} & \sigma_{31} & \sigma_{41} & \sigma_{51} & \sigma_{61} & \sigma_{71} \\ & \sigma_2^2 & \sigma_{32} & \sigma_{42} & \sigma_{52} & \sigma_{62} & \sigma_{72} \\ & & \sigma_3^2 & \sigma_{43} & \sigma_{53} & \sigma_{63} & \sigma_{73} \\ & & & \sigma_4^2 & \sigma_{54} & \sigma_{64} & \sigma_{74} \\ & & & & \sigma_5^2 & \sigma_{65} & \sigma_{75} \\ & & & & & \sigma_6^2 & \sigma_{76} \\ & & & & & & \sigma_7^2 \end{bmatrix} \quad (9)$$

2.3.8. Unstructured Correlation (UNC)

UNC structure is the most complex because it is estimating unique correlations for each pair of time points.

$$\text{UNC} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} & \sigma_{15} & \sigma_{16} & \sigma_{17} \\ & \sigma_2^2 & \sigma_{23} & \sigma_{24} & \sigma_{25} & \sigma_{26} & \sigma_{27} \\ & & \sigma_3^2 & \sigma_{34} & \sigma_{35} & \sigma_{36} & \sigma_{37} \\ & & & \sigma_4^2 & \sigma_{45} & \sigma_{46} & \sigma_{47} \\ & & & & \sigma_5^2 & \sigma_{56} & \sigma_{57} \\ & & & & & \sigma_6^2 & \sigma_{67} \\ & & & & & & \sigma_7^2 \end{bmatrix} \quad (10)$$

2.3.9. Huynh – Feldt (HF)

Huynh – Feldt is a circular matrix in which the covariance between any two elements is equal to the average of their variances minus a constant. Neither the variances nor the covariances are constant.

$$\begin{bmatrix} \sigma_1^2 & \frac{\sigma_1 + \sigma_2}{2} - \lambda & \frac{\sigma_1 + \sigma_3}{2} - \lambda & \frac{\sigma_1 + \sigma_4}{2} - \lambda & \frac{\sigma_1 + \sigma_5}{2} - \lambda & \frac{\sigma_1 + \sigma_6}{2} - \lambda & \frac{\sigma_1 + \sigma_7}{2} - \lambda \\ & \sigma_2^2 & \frac{\sigma_2 + \sigma_3}{2} - \lambda & \frac{\sigma_2 + \sigma_4}{2} - \lambda & \frac{\sigma_2 + \sigma_5}{2} - \lambda & \frac{\sigma_2 + \sigma_6}{2} - \lambda & \frac{\sigma_2 + \sigma_7}{2} - \lambda \\ & & \sigma_3^2 & \frac{\sigma_3 + \sigma_4}{2} - \lambda & \frac{\sigma_3 + \sigma_5}{2} - \lambda & \frac{\sigma_3 + \sigma_6}{2} - \lambda & \frac{\sigma_3 + \sigma_7}{2} - \lambda \\ & & & \sigma_4^2 & \frac{\sigma_4 + \sigma_5}{2} - \lambda & \frac{\sigma_4 + \sigma_6}{2} - \lambda & \frac{\sigma_4 + \sigma_7}{2} - \lambda \\ & & & & \sigma_5^2 & \frac{\sigma_5 + \sigma_6}{2} - \lambda & \frac{\sigma_5 + \sigma_7}{2} - \lambda \\ & & & & & \sigma_6^2 & \frac{\sigma_6 + \sigma_7}{2} - \lambda \\ & & & & & & \sigma_7^2 \end{bmatrix} \quad (11)$$

2.3.10. First Order Ante – Dependence (ANTE(1))

First order ante – dependence is the covariance structure with different variances in the diagonal and different covariances in

the off diagonal.

$$\text{ANTE (1)} = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_1 & \sigma_1\sigma_3\rho_1\rho_2 & \sigma_1\sigma_4\rho_1\rho_3 & \sigma_1\sigma_5\rho_1\rho_4 & \sigma_1\sigma_6\rho_1\rho_5 & \sigma_1\sigma_7\rho_1\rho_6 \\ & \sigma_2^2 & \sigma_2\sigma_3\rho_2 & \sigma_2\sigma_4\rho_2\rho_3 & \sigma_2\sigma_5\rho_2\rho_4 & \sigma_2\sigma_6\rho_2\rho_5 & \sigma_2\sigma_7\rho_2\rho_6 \\ & & \sigma_3^2 & \sigma_3\sigma_4\rho_3 & \sigma_3\sigma_5\rho_3\rho_4 & \sigma_3\sigma_6\rho_3\rho_5 & \sigma_3\sigma_7\rho_3\rho_6 \\ & & & \sigma_4^2 & \sigma_4\sigma_5\rho_4 & \sigma_4\sigma_6\rho_4\rho_5 & \sigma_4\sigma_7\rho_4\rho_6 \\ & & & & \sigma_5^2 & \sigma_5\sigma_6\rho_5 & \sigma_5\sigma_7\rho_5\rho_6 \\ & & & & & \sigma_6^2 & \sigma_6\sigma_7\rho_6 \\ & & & & & & \sigma_7^2 \end{bmatrix} \quad (12)$$

2.4. Modify Variance – Covariance Structure

2.4.1. First Order Auto – Regressive

AR (1) model is given as $(1-\phi L)Y_t = \varepsilon_t$, where ε_t is an independent white noise error term with zero mean and constant variance $(0, \sigma^2)$, and constant correlation (ρ) .

The density of ε_t is $p(\varepsilon_t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{\varepsilon_t^2}{\sigma^2}}$

To calculate $E[Y_{t-1}]$ by taking expectation of both sides, we obtain,

$$\begin{aligned} E[Y_t] &= \phi E[Y_{t-1}] + E[\varepsilon_t] \\ &= \phi E[Y_{t-1}], \text{ since } E[\varepsilon_t] = 0 \end{aligned} \quad (13)$$

We know,

$$\begin{aligned} Y_{t-1} &= \phi Y_{t-2} + \varepsilon_{t-1} \\ \therefore E[Y_{t-1}] &= \phi E[Y_{t-2}], \end{aligned} \quad (14)$$

So that $E[Y_t] = \phi^2 E[Y_{t-2}]$,

Now, we need to find $E[Y_{t-2}]$, clearly this process will never end. It is possible to break this infinite regress since by definition $E[Y_t] = E[Y_{t-1}]$. It then follows that $E[Y_t] = \phi E[Y_t] \Rightarrow (1-\phi)E[Y_t] = 0$

Assuming that $\phi \neq 1$. We conclude that Y_t is an AR(1) given $E[Y_t] = 0$.

Same process can be followed to calculate $\text{var}[Y_t]$ for an AR(1) from equation above, we have

$$\begin{aligned} \text{Var}[Y_t] &= \text{Var}[\phi Y_{t-1} + \varepsilon_t] \\ &= \underbrace{\phi^2 \text{Var}[Y_{t-1}]}_{\text{Var}[Y_t]} + \underbrace{2\phi \text{Cov}[\text{Var}[Y_{t-1}], \varepsilon_t]}_{=0} + \underbrace{\text{Var}[\varepsilon_t]}_{=\sigma^2} \end{aligned} \quad (15)$$

Hence,

$$\text{Var}[Y_t] = \text{Var}[Y_{t-1}]$$

and ε_t is i.i.d and hence uncorrelated with Y_{t-1} . Solving for $\text{Var}[Y_t]$, we obtain

$$\text{Var}[Y_t] = \frac{\sigma^2}{1-\phi^2} \quad (16)$$

2.4.2. Factor Analysis

Factor Analytic (FA) is prevalent in educational statistics as well as psychometrics [11], [12]. The general factor analytic form is given as

$$\Lambda\Lambda' + D \quad (17)$$

where Λ is a lower triangular matrix and D is a diagonal one. The elements of Λ are called factor loading and those of D specific variances. The portion of the total variance modelled in Λ represents variability attributes to some underlying common factors, whereas the remaining portions are unique. Aside from these interpretation, the FA structures provide another useful and distinct way of modelling heterogeneity. For the two FA structures, all column of Λ' except for the first are assumed to equal 0. These are known as first order FA structures because they only model one common factor. Higher order FA structure are also possible, although in this case constraint may be placed on a few of the parameters in order to fix the rotation [13]. The FA (1) structure is given below as;

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ \lambda_5 \\ \lambda_6 \\ \lambda_7 \end{bmatrix} + \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ \lambda_5 \\ \lambda_6 \\ \lambda_7 \end{bmatrix} \begin{bmatrix} \sigma_1^2 & & & & & & \\ & \sigma_2^2 & & & & & \\ & & \sigma_3^2 & & & & \\ & & & \sigma_4^2 & & & \\ & & & & \sigma_5^2 & & \\ & & & & & \sigma_6^2 & \\ & & & & & & \sigma_7^2 \end{bmatrix} \quad (18)$$

Let $x(p \times 1)$ be a random vector with mean μ and covariance matrix Σ . Then we say that the k-factor model holds for x, if x can be written in the Form

$$x = \Lambda f + u + \mu \quad (19)$$

where $\Lambda(p \times k)$ is a matrix of constants and $f(k \times 1)$ and $u(p \times 1)$ are random vectors. The elements of f are called *common* factors and the elements of u *specific* or *unique* factors. We shall suppose

$$E(f) = 0, V(f) = I.$$

$$E(u) = 0, C(u_i, u_j) = 0, i \neq j \text{ and } C(f, u) = 0.$$

Denote the covariance matrix of u by

$V(u) = \psi = \text{diag}(\psi_{11}, \dots, \psi_{pp})$. Thus, all of the factors are uncorrelated with one another and further the common factors are each standardized to have variance 1. It is sometimes convenient to suppose that f and u (and hence x) are multnormally distributed. Note that,

$$x_i = \sum_{j=1}^k \lambda_{ij} f_j + u_i + \mu_i, i=1, \dots, p \quad (20)$$

so that

$$\sigma_{ii} = \sum_{j=1}^k \lambda_{ij}^2 + \psi_{ii} \quad (21)$$

Thus, the variance of x can be split into two parts. First,

$$h_i^2 = \sum_{j=1}^k \lambda_{ij}^2$$

is called the *communality* and represents the variance of x_i , which is shared with the other variables via the common factors. In particular $\lambda_{ij}^2 = c(x_i, f_j)$ represents the extent to which x_i depends on the j^{th} common factor. On the other hand ψ_{ij} is called the *specific* or *unique variance* and is due to the unique factor u_i ; it explains the variability in x_i not shared with the other variables. The validity of the k -factor model can be expressed in terms of a simple condition on Σ . We get $\Sigma = \Lambda\Lambda' + \psi$. The converse also holds. If Σ can be decomposed, then the k -factor model holds for x . However, f and u are not uniquely determined by x . Re-scaling the variables of x is equivalent to letting $y = Cx$, where $C = \text{diag}(c_i)$. If the k -factor model holds for x with $\Lambda = \Lambda$, and $\psi = \psi$ then,

$$\begin{aligned} y &= C\Lambda f + Cu + C\mu \\ \text{and} \\ V(y) &= C\Sigma C' = C\Lambda\Lambda' C + C\psi C \end{aligned} \quad (22)$$

Thus the k -factor model also holds for y with factor loading matrix $\Lambda = C\Lambda$ and specific variances $\psi = C\psi C = \text{diag}(C_i^2 \psi_{ij})$. Note that the factor loading matrix for the scaled variables y is obtained by scaling the factor loading matrix of the original variables (multiply the i^{th} row of Λ_x , by C_i). A similar comment holds for the specific variances. In other words, factor analysis is unaffected by a re-scaling of the variables. If the k -factor model holds, then it also holds if the factors are rotated; that is, if G is a $(k \times k)$ orthogonal matrix, then x can also be written as

$$x = (\Lambda G)(G'f) + u + \mu \quad (23)$$

Since the random vector $G'f$ also satisfies the conditions above, we see that the k -factor model is valid with new factors $G'f$ and new factor loadings ΛG . Thus, we can also write $\Sigma = \Lambda\Lambda' + \psi = (\Lambda G)(\Lambda G)' + \psi$. In fact, for fixed ψ , this

rotation is the only indeterminacy in the decomposition of Σ in terms of Λ and Σ that is, if $\Sigma = \Lambda\Lambda' + \psi = \Lambda^* \Lambda^{*'} - \psi$, then $\Lambda = \Lambda^* G$ for some orthogonal matrix G . This indeterminacy in the definition of factor loadings is usually resolved by rotating the factor loadings to satisfy an arbitrary constraint such as $\Lambda' \psi^{-1} \Lambda$ is diagonal, or $\Lambda' D^{-1} \Lambda$ is diagonal. $D = \text{diag}(\sigma_{11}, \dots, \sigma_{pp})$ where in either case, the diagonal elements are written in decreasing order, say. Both constraints are scale invariant and, except for possible changes of sign of the columns, Λ is then in general completely determined by either constraint. Note that when the number of factors $k = 1$, the constraint is irrelevant. Also, if some of the ψ_{ii} equal 0, then the constraint cannot be used. It is of interest to compare the number of parameters in Σ when Σ is unconstrained, with the number of free parameters in the factor model. Let s denote the difference. At first sight Λ and ψ contain $pk + P$ free parameters. However, we introduce $4k(k-1)$ constraints. Since the number of distinct elements of Σ is $\frac{1}{2}p(p+1)$ we see that

$$s = \frac{1}{2}p(p+1) - \{pk + p - \frac{1}{2}k(k-1)\} = \frac{1}{2}(p-k)^2 - \frac{1}{2}(p+k) \quad (24)$$

Usually it will be the case that $s > 0$. Then s will represent the extent to which the factor model offers a simpler interpretation for the behavior of x than the alternative assumption that $V(x) = \Sigma$. If $s \geq 0$ and Λ and ψ are known, then Σ can be written in terms of Λ and ψ , subject to the constraint on Λ . We observe a data matrix X whose information is summarized by the sample mean \bar{x} and sample covariance matrix S . The location parameter is not of interest here, and we shall estimate it by $\mu = \bar{x}$. The interesting problem is how to estimate Λ and ψ and hence $\Sigma = \Lambda\Lambda' + \psi$ from S : that is, we wish to find estimates $\hat{\Lambda}$ and $\hat{\psi}$ satisfying the constraint $s = \hat{\Lambda}\hat{\Lambda}' + \hat{\psi}$, at least approximately. Given an estimate $\hat{\Lambda}$, it is then natural to set

$$\hat{\psi}_{ii} = s_{ii} - \sum_{j=1}^k \hat{\lambda}_{ij}^2, i=1, \dots, p \quad (25)$$

So that the diagonal equations always hold exactly. We shall only consider estimates for which it is satisfied and $\hat{\psi}_{ii} > 0$. Setting $\hat{\Sigma} = \hat{\Lambda}\hat{\Lambda}' + \hat{\psi}$, we get $\hat{\sigma}_{ii} = \sum_{j=1}^k \hat{\lambda}_{ij}^2 + \hat{\psi}_{ii}$ is equivalent to the condition

$$\hat{\sigma}_{ii} = s_{ii}$$

Three cases can occur depending on the value of s . If $s < 0$ then it contains more parameters than above. Then, in general, we expect to find an infinity of exact solutions for

Λ and Ψ , and hence, the factor model is not well - defined. If $s = 0$ then it can generally be solved for Λ and Ψ exactly (subject to the constraint on Λ). The factor model contains as many parameters as Σ and hence offers no simplification of the original assumption that $V(x) = \Sigma$. However, the change in viewpoint can sometimes be very helpful. If $s > 0$, as will usually be the case, then there will be more equations than parameters. Thus, it is not possible to solve exactly in terms of $\hat{\Lambda}$ and $\hat{\Psi}$, we must look for approximate solutions. In this case the factor model offers a simpler explanation for the behaviour of x than the full covariance matrix. Because the factor model is scale invariant, we shall only consider estimates of $\Lambda = \hat{\Lambda}$, and $\Psi = \hat{\Psi}$, which are scale invariant. It is then convenient to consider the scaling separately from the relationships between the variables. Let $Y = HXD_s^{-\frac{1}{2}}$ where $D_s = \text{diag}(s_{11}, \dots, s_{pp})$, denote the standardized variables so that

$$\sum_{r=1}^n y_{rj} = 0, \text{ and } \frac{1}{n} \sum_{r=1}^n y_{rj}^2 = 1, j = 1, \dots, p$$

Then Y will have estimated factor loading matrix $\hat{\Lambda}_y = D_s^{-\frac{1}{2}} \hat{\Lambda}_x$ and estimated specific variances

$$\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\lambda_1\lambda_2 & \sigma_1\sigma_3\lambda_1\lambda_3 & \sigma_1\sigma_4\lambda_1\lambda_4 & \sigma_1\sigma_5\lambda_1\lambda_5 & \sigma_1\sigma_6\lambda_1\lambda_6 & \sigma_1\sigma_7\lambda_1\lambda_7 \\ & \sigma_2^2 & \sigma_2\sigma_3\lambda_2\lambda_3 & \sigma_2\sigma_4\lambda_2\lambda_4 & \sigma_2\sigma_5\lambda_2\lambda_5 & \sigma_2\sigma_6\lambda_2\lambda_6 & \sigma_2\sigma_7\lambda_2\lambda_7 \\ & & \sigma_3^2 & \sigma_3\sigma_4\lambda_3\lambda_4 & \sigma_3\sigma_5\lambda_3\lambda_5 & \sigma_3\sigma_6\lambda_3\lambda_6 & \sigma_3\sigma_7\lambda_3\lambda_7 \\ & & & \sigma_4^2 & \sigma_4\sigma_5\lambda_4\lambda_5 & \sigma_4\sigma_6\lambda_4\lambda_6 & \sigma_4\sigma_7\lambda_4\lambda_7 \\ & & & & \sigma_5^2 & \sigma_5\sigma_6\lambda_5\lambda_6 & \sigma_5\sigma_7\lambda_5\lambda_7 \\ & & & & & \sigma_6^2 & \sigma_6\sigma_7\lambda_6\lambda_7 \\ & & & & & & \sigma_7^2 \end{bmatrix} \quad (26)$$

The unique solution of the sign $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6$, and λ_7 is given as $\lambda_1 = \sqrt{r_{12}r_{13}/r_{23}}$, $\lambda_2 = \sqrt{r_{12}r_{23}/r_{13}}$, $\lambda_3 = \sqrt{r_{13}r_{23}/r_{12}}$, e.t.c

2.5. Goodness of Fit Criteria

Information criteria are used when comparing or examining different models for the same data. Akaike's information criterion (AIC), Burham – Anderson criterion (AICC), Bayesian information criterion (BIC), Bozdogan Information Criterion (CAIC), Hannan – Quinn Information Criterion (HQIC) and the modified Adjusted Schwarz Information Criterion (ASIC) shall be used in this research to examine the best covariance structure for repeated measures. The smaller the fit criterion, the better [14]. The test for goodness of fit determines whether a set of observed data conforms to a specified probability distribution.

$\hat{\Psi}_y = D_s^{-\frac{1}{2}} \hat{\Psi}_x$ and can be written in terms of the correlation matrix of x as $R = \hat{\Lambda}_y \hat{\Lambda}_y' + \hat{\Psi}_y$. Note that

$$\hat{\Psi}_{yii} = 1 - \sum_{j=1}^k \hat{\lambda}_{yij}^2, i = 1, \dots, p$$

so that Ψ_y is not a parameter of the model any more, but a function of Λ . However, R contains p fewer free parameters than S so that s , the difference between the number of equations and the number of free parameters is still given by the p equations for the estimates of the scaling parameters are given. Since in practice it is the relationship between the variables which is of interest rather than their scaling, the data is often summarized by R rather than S . The scaling estimates are then not mentioned explicitly, and the estimated factor loadings and specific variances are presented in terms of the standardized variables.

2.4.3. First Order Auto Regressive Factor Analysis (ARFA(1))

First Order Auto – Regressive Factor Analysis ARFA (1). The combine model for ARFA (1) variance – covariance structure is a heterogeneous structure which neither the diagonal nor off diagonal are the same, is given as $(1 - \phi_L) \Lambda \Lambda' + (1 - \phi_L) D$.

2.6. Adjusted Goodness of Fit Criterion

Looking at Bayesian Information Criterion and adjust it, using the model below

$$f(k) = \{f(y/\theta_k) / \theta_k \in \Theta(k)\} \quad (27)$$

Fitted model is given as

$$f(y/\hat{\theta}_k)$$

$$\text{BIC} = -2 \ln f(y/\hat{\theta}_k) + 2k \ln n \quad (28)$$

We will consider a simple adjustment or justification of BIC in this research

Let y denote the observed data, assume that y is to be described using a model M_k selected from a set of candidate

models $M_{k1}, M_{k2}, \dots, M_{kl}$. Also, assume that each M_k is uniquely parameterized by a vector θ_k , where θ_k is an element of the parameter space Θ_k ($k \in \{k_1, k_2, \dots, k_l\}$).

Let $L(\theta_k / y)$ denote the likelihood for y based on M_k , where $L(\theta_k / y) = f(y / \theta_k)$

Let $\hat{\theta}_k$ denote the maximum likelihood estimate of θ_k obtained by maximizing $L(\theta_k / y)$ over Θ_k

Assume that derivatives of $L(\theta_k / y)$ up to order two exist with respect to θ_k and are continuous and suitably bounded for all $\theta_k \in \Theta(k)$. Let $\pi(k) (k \in \{k_1, k_2, \dots, k_l\})$

denote a discrete prior over the models $M_{k1}, M_{k2}, \dots, M_{kl}$. (θ_k / k) denote a prior on θ_k given model M_k ($k \in \{k_1, k_2, \dots, k_l\}$).

From the above, we are going to applying Bayes' theorem, the joint posterior of M_k and θ_k can be written as

$$h((k, \theta_k) / y) = \frac{\pi(k) g(\theta_k / k) L(\theta_k / y)}{m(y)} \quad (29)$$

$m(y)$ denotes the marginal distribution of (y) .

The posteriori probability for M_k is given by

$$p(k / y) = m(y)^{-1} \pi(k) \int L(\theta_k / y) g(\theta_k / k) d\theta_k. \quad (30)$$

Consider minimizing $-2\ln P(k/y)$ as opposed to maximizing $P(k/y)$.

We have

$$-2\ln p(k / y) = 2\ln\{m(y)\} - 2\ln\{\pi(k)\} - 2\ln\left\{\int L(\theta_k / y) g(\theta_k / k) d\theta_k\right\} \quad (31)$$

$m(y)$ is constant with respect to k ; thus, for the purpose of model selection, the term can be discarded. Hence, we obtain

$$\begin{aligned} & -2\ln p(k / y) \propto -2\ln\{\pi(k)\} - 2\ln\left\{\int L(\theta_k / k) d\theta_k\right\} \\ & \equiv s(k / y) \end{aligned} \quad (32)$$

We can use the integral above in order to obtain the approximation of the above term by taking a second order Taylor series expansion of the log-likelihood about $\hat{\theta}_k$ as

$$\begin{aligned} \ln L(\theta_k / y) & \approx \ln L(\hat{\theta}_k / y) + (\theta_k - \hat{\theta}_k)' \frac{\partial \ln L(\hat{\theta}_k / y)}{\partial \theta_k} + \frac{1}{2} (\theta_k - \hat{\theta}_k)' \left[\frac{\partial^2 \ln L(\hat{\theta}_k / y)}{\partial \theta_k \partial \theta_k'} \right] (\theta_k - \hat{\theta}_k) \\ & = \ln L(\hat{\theta}_k / y) - \frac{1}{2} (\theta_k - \hat{\theta}_k)' \left[n \bar{I}(\hat{\theta}_k, y) \right] (\theta_k - \hat{\theta}_k) \end{aligned} \quad (33)$$

where

$$\bar{I}(\hat{\theta}_k, y) = -\frac{1}{n} \frac{\partial^2 \ln L(\hat{\theta}_k / y)}{\partial \theta_k \partial \theta_k'}$$

is the average observed fisher information matrix

Thus,

$$L(\theta_k / y) \approx L(\hat{\theta}_k / y) \exp \left\{ -\frac{1}{2} (\theta_k - \hat{\theta}_k)' \left[n \bar{I}(\hat{\theta}_k, y) \right] (\theta_k - \hat{\theta}_k) \right\} \quad (34)$$

We therefore have the following approximation for our integral as

$$\int L(\theta_k / y) g(\theta_k / k) d\theta_k \approx L(\hat{\theta}_k / y) \int \exp \left\{ -\frac{1}{2} (\theta_k - \hat{\theta}_k)' \left[n \bar{I}(\hat{\theta}_k, y) \right] (\theta_k - \hat{\theta}_k) \right\} g(\theta_k / k) d\theta_k \quad (35)$$

Consider the evaluation of

$\int \exp \left\{ -\frac{1}{2} (\theta_k - \hat{\theta}_k)' \left[n \bar{I}(\hat{\theta}_k, y) \right] \right\} g(\theta_k / k) d\theta_k$ using the non-informative prior $g(\theta_k / k) = 1$

We obtain

$$\int \exp \left\{ -\frac{1}{2} (\theta_k - \hat{\theta}_k)' \left[n \bar{I}(\hat{\theta}_k, y) \right] \right\} g(\theta_k / k) d\theta_k = (2\pi)^{(k/2)} \left| n \bar{I}(\hat{\theta}_k, y) \right|^{-\frac{1}{2}} \quad (36)$$

We therefore have

$$\begin{aligned} \int L(\theta_k / y) g(\theta_k / k) d\theta_k &\approx L(\hat{\theta}_k / y) (2\pi)^{(k/2)} \left| n \bar{I}(\hat{\theta}_k, y) \right|^{-\frac{1}{2}} \\ &= L(\hat{\theta}_k / y) (2\pi)^{(k/2)} n^{(-k/2)} \left| \bar{I}(\hat{\theta}_k, y) \right|^{-\frac{1}{2}} \\ &= L(\hat{\theta}_k / y) \left(\frac{2\pi}{n} \right)^{(k/2)} \left| \bar{I}(\hat{\theta}_k, y) \right|^{-\frac{1}{2}} \end{aligned} \quad (37)$$

The preceding can be viewed as a variation on the laplace method of approximating the integral

$$\int L(\theta_k / y) g(\theta_k / k) d\theta_k.$$

Then, going back to our $S(k/y)$, we can write

$$\begin{aligned} s(k / y) &= -2 \ln \{ \pi(k) \} \\ &= -2 \ln \left\{ \int L(\theta_k / y) g(\theta_k / k) d\theta_k \right\} \\ &\approx -2 \ln \{ \pi(k) \} \\ &= -2 \ln \left[L(\hat{\theta}_k / y) \left(\frac{2\pi}{n} \right)^{(k/2)} \left| \bar{I}(\hat{\theta}_k, y) \right|^{-\frac{1}{2}} \right] \\ &= -2 \ln \{ \pi(k) \} - 2 \ln L(\hat{\theta}_k / y) + k \left\{ \ln^2 \left(\frac{n}{2\pi} \right) \right\} + \ln^2 \left| \bar{I}(\hat{\theta}_k, y) \right| \end{aligned} \quad (38)$$

Ignoring terms $\ln^2 \left| \bar{I}(\hat{\theta}_k, y) \right|$ that is bounded as the sample size grows to infinity, we obtain our adjusted model as

$$\begin{aligned} s(k / y) &\approx -2 \ln L(\hat{\theta}_k / y) + k \ln \ln(n) \\ ASIC &= -2 \ln L(\hat{\theta}_k / y) + k \ln \ln(n) \\ ASIC &= -2 \ln L(y / \hat{\theta}_k) + k \ln \ln(n) \end{aligned} \quad (39)$$

3. Analysis

Table 1. Mauchly's Test of Sphericity.

Within Subject Effects	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon ^b		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
WEEKS	.004	440.232	20	.000	.493	.544	.167

Table 1 shows that the assumption of Sphericity has not been met, since the Sig. value is 0.00 which is less than 0.05, so the null hypothesis that the variances of the difference between levels were significantly the same was rejected.

Table 2. Fitting Criteria Results for Covariance Structures Covariance Structure.

Information Criteria	UN	CS	HF	AR (1)	ARH (1)
AIC	1560.159	1609.557	1597.026	1609.557	1487.235
AICC	1581.812	1609.676	1598.541	1609.676	1488.750
BIC	1634.202	1614.846	1618.181	1614.846	1508.390
HQIC	1548.753	1610.170	1594.866	1610.170	1485.075
CAIC	1662.202	1616.846	1626.181	1616.846	1516.390

Covariance Structures Cont.

Information Criteria	CSH	TOEP	TOEPH	ANTE(1)	UNC
AIC	1621.557	1619.557	1497.236	1497.235	1527.235
AICC	1623.073	1620.723	1501.280	1501.279	1548.888
BIC	1642.712	1638.067	1531.613	1531.612	1601.277
HQIC	1619.397	1617.859	1492.764	1492.763	1515.829
CAIC	1650.390	1645.067	1544.613	1544.612	1629.277

Table 2 shows the Information criteria for ten different covariance structures in mixed model approach. According to AIC, AICC, BIC, HQIC and CAIC fitting criteria, Heterogeneous first order autoregressive ARH (1) was the best covariance structure (since it is the one that provide the most smallest value on both the HQIC, AIC and AICC) then followed by first order ante-dependence ANTE(1), then Heterogeneous toeplitz (TOEPH), then unstructured correlation (UNC), then unstructured (UN), followed by Huynh-Feldt (HF), then followed by compound symmetry (CS) and first order auto-regressive AR(1), then toeplitz (TOEP) while the worst one was Heterogeneous compound symmetry (CSH). ARH (1) and ANTE(1) gave informations about growth-development mechanism and consecutive variation at weight performances of the broilers in grams over trial time. This indicates that Heterogeneous first order autoregressive and first order ante dependence are the best

covariance structures for this research and was selected for further examination with the modified structures and information criteria still in this research.

Table 3. Information Criteria for ARFA(1).

ASIC	HQIC	AIC	AICC	BIC	CAIC
1131.094	1131.095	1137.567	1138.967	1174.722	1151.864

Table 3 shows the result of first order auto regressive factor analysis ARFA(1) using six different information criteria. From the result above it indicates that adjusted Schwarz information criteria (ASIC) was the best information criteria, because it gives the smallest value as 1131.094, then followed by Hannan – Quinn information criteria (HQIC) as 1131.095, then akaike information criteria (AIC) as 1137.567 while Bayesian information criteria (BIC) as 1174.722 is the worst.

Table 4. Summary of Covariance Structures with Best two Information Criteria.

COVARIANCE STRUCTURES	No. of p	n	RESTRICTED LOG LIKELIHOOD	HQIC	ASIC
UN	28	105	1504.159	1548.753	1547.215
CS	2	105	1605.557	1610.17	1608.632
HF	8	105	1581.026	1594.866	1593.328
AR(1)	2	105	1605.557	1610.17	1608.632
ARH(1)	8	105	1471.235	1485.075	1483.537
CSH	8	105	1605.557	1619.397	1617.859
TOEP	7	105	1605.557	1617.859	1616.321
TOEPH	13	105	1471.235	1492.763	1491.226
ANTE(1)	13	105	1471.235	1492.763	1491.225
UNC	28	105	1471.235	1515.829	1514.291
ARFA(1)	14	105	1109.567	1131.095	1131.094

Table 4 shows the summary of eleven covariance structures, each attached with the parameter numbers using seven repeated measure data. It was clearly indicates that ARFA(1) was the best covariance structures, then followed by ARH(1), then followed by ANTE(1) and TOEPH, while CSH was the worst in this research. This happen as the result of violation of sphericity assumption. Also in terms of information criteria ASIC gives the lowest values throughout, such it was the best information criteria for comparison covariance structures.

4. Conclusion

In this research, The sphericity assumption was violated according to mauchly statistic result ($P < 0.05$) and according to information criteria AIC, AICC, BIC, HQIC and CAIC Heterogeneous first order autoregressive ARH (1) was the best covariance structure (since it is the one that provide the most smaller value on both the HQIC, AIC, BIC, CAIC and AICC) then followed by first order ante-dependence

ANTE(1), then Heterogeneous toeplitz (TOEPH), then unstructured correlation (UNC), then unstructured (UN), followed by Huynh-Feldt (HF), then followed by compound symmetry (CS) and first order auto-regressive AR(1), then toeplitz (TOEP) while the worst one was Heterogeneous compound symmetry (CSH). ARH (1) and ANTE(1) give information about growth-development mechanism and consecutive variation at weight performances of the broilers in grams over trial time. This indicates that Heterogeneous first order autoregressive and first order ante dependence are the best covariance structures for this research and was selected for further examination with the modified structure also in this research as the best comparison structures using information criteria. Hence when examining it with the modified structure ARFA(1) was the best structure while ASIC gives the smallest values throughout, then followed by HQIC. The overall result indicates ARFA(1) as the best covariance structure, followed by ARH(1) while CSH was the worst structure in this research and ASIC gives the smallest values, followed by HQIC as the best information criteria for examination of covariance structures. SPSS version 23.6 and R version 3.4.0 was used for the analysis.

References

- [1] J. Algina, R. R. Wilcox and R. K. Kowalchuk, "The Analysis of Repeated Measure. A Quantitative Research Synthetic," *British Journal of Mathematical and Statistical Psychology*, pp. 1735-1748, 2000.
- [2] N. K. Rauf, "Classification of Multivariate Repeated Measures data with Temporal Autocorrelation," *Advances in Data Analysis and Classification*, pp. 175-199, 2008.
- [3] L. A. Guttman, New Approach to Factor Analysis: The Radex in P. F. Lazarsfeld (Ed.), *Mathematical thinking in the Social sciences*, New York: Columbia University Press, 1954.
- [4] L. Brandon, Misspecification of the Covariance Matrix in the Linear Mixed Model: A Monte Carlo Simulation, MN, USA: University of Minnesota, 2013.
- [5] D. J. Yanosky II, Comparability of Covariance Structures and Accuracy of Information Criteria in Mixed Model Methods for Longitudinal Data Analysis., Athens, Georgia: University of Georgia Press, 2007.
- [6] J. Ferron, R. Dailey and Q. Yi, "Effect of Misspecifying the First Level error Structure in Two Level Models of Change," *Multivariate Behav. Res.*, pp. 379-403, 2002.
- [7] S. D. Maxwell and H. D. Delaney, *Designing Experiments and Analyzing Data: A Model Comparison Perspective*, 2nd ed., Mahwah, NJ: Laurence Erlbaum Associated Publishers, 2004.
- [8] E. R. Girden, *ANOVA: Repeated Measures*, Sage: Newbury Park CA, 1992.
- [9] S. W. Greenhouse and S. Geisser, "On Methods in the Analysis of Profile Data," *Psychometrika*, pp. 95-112, 1959.
- [10] H. Huynh and L. S. Feldt, "Estimation of the Box Correction for Degrees of Freedom from Sample Data in the Randomized Block and Split Plot Design," *Journal of Educational Statistics*, pp. 15-51, 1976.
- [11] R. P. McDonald, *Factor Analysis and Related Methods*, Lawrence Erlbaum Associates, 1985.
- [12] S. A. Mulaik, *The Foundation of Factor Analysis*, New York: McGraw Hill, 1972.
- [13] R. I. Jennrich and M. D. Schluchter, "Unbalanced Repeated Measures Models with Structured Covariance Matrix," *Biometrics*, pp. 805-820, 1986.
- [14] R. C. Littell, G. A. Milleken, W. W. Stroup and R. D. Wolfinger, *SAS System for Mixed Models*, Cary: NC; SAS Institute, 1996.