A STUDY OF COVARIANCE STRUCTURE SELECTION FOR SPLIT-PLOT DESIGNS
ANALYZED USING MIXED MODELS

by

CHEN QIU

B.S., Shanghai University of Traditional Chinese Medicine, 2008

A REPORT

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Statistics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2014

Approved by:

Major Professor
Dr. Christopher I. Vahl

Copyright

CHEN QIU

2014

Abstract

In the classic split-plot design where whole plots have a completely randomized design, the conventional analysis approach assumes a compound symmetry (CS) covariance structure for the errors of observation. However, often this assumption may not be true. In this report, we examine using different covariance models in PROC MIXED in the SAS system, which are widely used in the repeated measures analysis, to model the covariance structure in the split-plot data in which the simple compound symmetry assumption does not hold. The comparison of the covariance structure models in PROC MIXED and the conventional split-plot model is illustrated through a simulation study. In the example analyzed, the heterogeneous compound symmetry (CSH) covariance model has the smallest values for the Akaike and Schwarz's Bayesian information criteria fit statistics and is therefore the best model to fit our example data.

Keywords: split-plot, covariance structure, repeated measures, mixed model

Table of Contents

List of Tables

Acknowledgements

Firstly, I would like to express my sincere thanks to my major professor Dr. Christopher I. Vahl for his instruction, support and encouragement during the time of my master's research.

Then, I would also like to express my appreciation to my committee members Dr. Leigh Murray and Dr. Weixin Yao for all the time and efforts they have spent on my research.

Especially, I greatly appreciate Dr. Leigh Murray for teaching me how to behave as a good statistical consultant.

Also I would like to express my gratitude to Dr. Alexander M. Stelzleni from University of Georgia for generously sharing his research data with me.

I extend my sincere gratitude to Dr. Guihua Bai for providing me with the GRA position.

Final, my special thanks go to my husband, Dr. Lu, for his love and care, and also to my little sweetheart, Rachel, for all the incredible joy she has brought and will continue to bring to my life.

CHAPTER 1 - Introduction

1.1 Split-plot in Completely Randomized Design (CRD)

Split-plot designs were first introduced by Fisher (1925) in the agriculture experiments and widely used in the industrial experiments (Kowalski and Potcner, 2003) for the reason that some treatment is hard to change, either practically or economically.

A split-plot experiment is regarded as a blocked experiment (Mbegbu and Francis, 2012), where the hard-to-change factors are applied to the blocks, which are referred to as the whole plots, and the other factors are then applied to the small experimental units, which are referred to as the subplots or split-plots. Therefore, there are two levels of experimental units in the split-plot design. Corresponding to the two levels of experimental units, there are two levels of randomization. One randomization is conducted to assign the whole plot treatment to the whole plot experimental units, and the other is conducted to assign the subplot treatment to the subplot experimental units. Owing to the two levels of randomization, there are two sources of errors in the split plot model.

Split-plot designs are considered for use when a completely randomized design is structurally impossible, expensive, inefficiency or less valid, see Jones and Nachsheim (2009) for details. However, due to the existence of two sources of errors, the analysis of a split-plot design is more complicated than that for a completely randomized design. The usual assumptions made about the errors in split-plot designs are that whole-plot errors are independently and identically distributed (iid) as $N(0, \sigma_w^2)$, subplot errors are iid $N(0, \sigma_e^2)$, and the whole-plot errors and subplot errors are independent.

The ANOVA model (for example, see Box and Jones, 1990) for the response in a balanced two-factor split-plot design where whole plots are in a CRD with a levels of whole plot factor A, b levels of subplot factor B, r replications of the ith level of factor A, and the number of observations $n = abr$, is given by:

$$y_{ijk} = \mu + \alpha_i + w_{ik} + \beta_j + (\alpha\beta)_{ij} + e_{ijk} \tag{1.1}$$

Where μ is a constant and $\mu_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$ is the mean of the ij^{th} A by B treatment combination with α_i as the effect of the ith level of whole-plot factor A, β_j as the effect of the jth

level subplot factor B, and $(\alpha\beta)_{ij}$ as the interaction effect of the ith level of A with the jth level of B (or the A by B interaction effect). Here w_{ik} is the random error associated with the kth whole plot assigned to the ith level of A and e_{ijk} is the random error associated with the subplot assigned to the jth level of B within the kth whole plot assigned to the ith level of A. As noted above, it is assumed that $w_{ik} \sim \text{iid } N(0,\sigma_w^2)$, $e_{ijk} \sim \text{iid } N(0,\sigma_e^2)$, and all w_{ik} and e_{ijk} terms are assumed to be independent.

The model given in Equation 1.1 can be written in matrix form as

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Zu} + \mathbf{e}$$

Where $\mathbf{y}$ is the $n \times 1$ vector of responses; $\mathbf{X}$ is the $n \times p$ design matrix for the fixed effects parameters, $\mathbf{\beta}$ is the p-dimensional vector of unknown fixed effects parameters; $\mathbf{Z}$ is the $n \times ar$ design matrix for the whole-plot errors; $\mathbf{u}$ is the ar-dimensional vector of unknown random whole plot errors; and $\mathbf{e}$ is the n-dimensional vector of unknown random subplot errors. In this model it is assumed that $\mathbf{u}$ is a multivariate normal random vector with a mean vector containing all zeroes and a variance-covariance matrix that has σ_w^2 for all diagonal elements and all zeroes for the off-diagonal elements, i.e. $\mathbf{u} \sim N(\mathbf{0}, \sigma_w^2 \mathbf{I}_{ar})$ where $\mathbf{0}$ is a vector/matrix of zeroes of appropriate dimension and $\mathbf{I}_{ar}$ is an $ar \times ar$ identity matrix. Similarly, we assume that $\mathbf{e} \sim N(\mathbf{0}, \sigma_e^2 \mathbf{I}_n)$ where vectors $\mathbf{u}$ and $\mathbf{e}$ are independent. Based on the above model assumptions, the covariance matrix of the response vector $\mathbf{y}$ can be expressed as

$$\mathbf{\Sigma} = \sigma_w^2 \mathbf{ZZ'} + \sigma_e^2 \mathbf{I}_n$$

where $\mathbf{Z'}$ denotes the transpose of matrix $\mathbf{Z}$. If the observations are grouped by whole plots, the covariance matrix of the response vector is the block-diagonal matrix given by

$$\mathbf{\Sigma} = diag\left(\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_{ar}\right)$$

where the ith element on the diagonal is the same $b \times b$ matrix $\mathbf{V}$ with the form of

$$\mathbf{V} = \sigma_e^2 \mathbf{I}_b + \sigma_w^2 \mathbf{J}_b \tag{1.2}$$

or simplified as

$$\mathbf{V} = \sigma_e^2 (\mathbf{I}_b + \eta \mathbf{J}_b)$$

where $\mathbf{J}_b$ is a $b \times b$ matrix of ones and $\eta = \sigma_w^2 / \sigma_e^2$ measures how much the subplots within the same whole plot are correlated (*cf*. Vahl and Milliken, 2011). In a more general form, the mixed model can be written as $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}^*$, where $\mathbf{X}$ and $\boldsymbol{\beta}$ as before but now $\mathbf{e}^* \sim N(\mathbf{0}, \boldsymbol{\Sigma})$.

From Equation (1.2) we can note two prominent features of $\mathbf{V}$: (1) the same value of $\sigma_e^2 + \sigma_w^2$ appears on the diagonal and (2) the same value of σ_w^2 appears on the off-diagonal. Feature (1) implies that all subplot treatment factors have the same variance and Feature (2) implies the covariance between any two subplots within the same whole plot is σ_w^2 which, because it is a variance, must always be non-negative. This variance-covariance structure obtained through the standard split-plot analysis is identical to fitting a repeated measures structure to the subplots assuming compound symmetry. For a repeated measures model with compound symmetry, the variance covariance matrix within a whole plot or subject (in the language of repeated measures) is given by

$$\mathbf{V} = \sigma^2 \begin{bmatrix} 1 & \rho & \rho & \cdots & \rho \\ \rho & 1 & \rho & \cdots & \rho \\ \rho & \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \rho & \cdots & 1 \end{bmatrix}. \tag{1.3}$$

If we let $\sigma^2 = \sigma_e^2 + \sigma_w^2$ and $\rho = \sigma_w^2 / (\sigma_e^2 + \sigma_w^2)$, note that $\rho = \eta / (\eta + 1)$, we see that the variance-covariance structure obtained through the usual split-plot analysis is a special case of compound symmetry where the correlation between any two subplots within the same whole plot is non-negative. Note that compound symmetry does not require that $\rho \geq 0$, only that $-1 \leq \rho \leq 1$ (*cf*. Littell *et al.*, 1996).

However, the randomization of subplot factor levels to subplots within whole plots does not guarantee the whole plot errors will be mutually independent unless the application of treatments to subplots is also applied completely independently. If this is not the case, the split-plot errors may not be mutually independent within whole plots, and so the assumption of the compound symmetry (CS) may not be justified. For example, consider an experiment to study the effect of temperature and baking time on bread. The researcher bakes three loaves of bread in an oven at the same time and, after 8 minutes, he takes one randomly selected loaf of bread out and measures a response. After 10 minutes, he takes out another at random and measures its

response. Finally, after 12 minutes, he takes out the last loaf of bread and measures its response. Here, the levels of temperature are assigned to all three loaves of bread baked at the same time, but a baking time is randomly assigned to each loaf in the oven resulting in a split-plot in randomized complete block design. Obviously, this is not a repeated measures design since the same unit (i.e. a loaf) was not measured three different times. However, this kind of split-plot experiment is like a repeated measures experiment in that, although baking time was randomly assigned, the same application of baking temperature overlaps for the three loaves. It would be expected that two loaves baked closer together in time would be more similar than those further apart. In other words, it is reasonable to conclude that the loaves been taken out after 8 and 10 minutes are more similar than the loaves taken out after 8 and 12 minutes. There also may be some drift in the response as time elapses causing an increase (or decrease) in the variance. Thus, for this kind of split-plot experiment, the conventional analysis approach which requires the assumption of a CS covariance structure to be hold, may not work very well.

Because of the abovementioned problem, we look for other approaches that do not require the covariance structure to be a particular form, and this leads us to consider other variance-covariance structures in the mixed model.

1.2 Mixed Models

In the last a few decades, mixed models have been widely used in fitting data from agronomy, social science, pharmaceuticals industry, economics, physical science and etc. (Wolfinger, Tobias and Sall, 1991). The increase of use of mixed models is partly due to the development of mixed models software, such as the MIXED procedure (PROC MIXED) in the SAS system and JMP software (Littell, Milliken, Stroup, and Wolfinger, 1996). The mixed model is particularly useful in the analysis of animal experiments with repeated measures data (Wang and Goonewardene, 2004), because it embeds the structure and relationships among the errors, and thus is more flexible than the classic repeated measures approaches which assume a very simple structure of dependence among errors.

1.3 Mixed Models Analysis of Data from Repeated Measures Design

The repeated measures design refers to multiple, or repeated, measurements made on the same experimental unit or subject which is observed over time. A repeated measures study usually consists of a completely randomized design with data been collected sequentially over

time (Littell *et al*., 2006). In a repeated measures design, often, measurements taken on the same subject are more likely to be correlated than those taken on different subjects, and two measurements taken close together on the same subject are likely to be more correlated than those taken further apart. Therefore, it is critical to properly identify the covariance structure of the errors in the analysis of the repeated measures (Littell, Pendergast, and Natarajan, 2000).

Due to the characteristics of repeated measures data, the conventional univariate ANOVA and multivariate ANOVA repeated approaches often work poorly compared to the mixed model approach using PROC MIXED in the SAS system (Wang and Goonewardene, 2004). The univariate ANOVA approach assumes a compound symmetric covariance structure of errors, and it cannot deal with the problem of unequal time spacing or missing data (Wang and Goonewardene, 2004). The multivariate ANOVA approach assumes only an unstructured covariance structure and cannot directly describe the covariance structure of errors, and it also requires data to be balanced (Wang and Goonewardene, 2004). While the mixed model approach in the SAS PROC MIXED can take the covariance structure of errors into consideration, and also it has the ability to handle missing data and unequal time spacing.

The covariance structure of repeated measures often can be easily obtained by using the REPEATED statement in PROC MIXED in SAS system and the "TPYE =" option in the REPEATED statement determines the covariance structure of repeated measures (Littell *et al.*, 2006).

1.4 Review of Some Candidate Covariance Structures in Mixed Model Procedure

1.4.1 Mixed Model Analysis with AR(1) Covariance Structure

The first-order autoregressive model {AR(1)} is a classic model in the time series analysis, and it is based on the idea that the current value of the process Y_t depends on its past value Y_{t-1}. The covariance structure of AR(1) assumes the correlation between any two observations that are right next to each other, such as 1st and 2nd, 2nd and 3rd, 3rd and 4th, and so on, is ρ; and the correlation between any two observations that have been separated by $n-1$ other observations, is ρ^n. And it also assumes homogeneous variances among all the observations. The variance-covariance matrix of AR(1) is expressed as following:

$$\mathbf{V} = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{k-1} \\ \rho & 1 & \rho & \cdots & \rho^{k-2} \\ \rho^2 & \rho & 1 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \rho \\ \rho^{k-1} & \rho^{k-2} & \cdots & \rho & 1 \end{bmatrix}$$

where k is the number of repeated measurements per experimental unit. From the matrix form of AR(1) covariance structure above, we can see that, if $-1 < \rho < 1$, two observations tend to get less and less correlated as they get farther and farther apart. The variances across the time points are all equal because the same value of σ^2 is on the main diagonal of the matrix. Also note the AR(1) model requires only two parameters: σ^2 and ρ.

1.4.2 Mixed Model Analysis with ARH(1) Covariance Structure

Unlike the AR(1) covariance structure which assumes homogeneous variances among all the observations, the heterogeneous first-order autoregressive {ARH(1)} allows the variances to be different. The variance–covariance matrix of ARH(1) is expressed as following:

$$\mathbf{V} = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho^2 & \cdots & \sigma_1\sigma_k\rho^{k-1} \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \cdots & \sigma_2\sigma_k\rho^{k-2} \\ \sigma_3\sigma_1\rho^2 & \sigma_3\sigma_2\rho & \sigma_3^2 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \sigma_{k-2}\sigma_k\rho \\ \sigma_k\sigma_1\rho^{k-1} & \sigma_k\sigma_2\rho^{k-2} & \cdots & \sigma_k\sigma_{k-2}\rho & \sigma_k^2 \end{bmatrix}$$

From the matrix form of ARH(1) covariance structure above, we can see that the variances of the main diagonal of the matrix are different. Therefore, the ARH(1) models requires more parameters than AR(1) model. The ARH(1) model requires $k+1$ covariance parameters, where k represents the number of the observations per subject.

1.4.3 Mixed Model Analysis with CSH Covariance Structure

The only difference between CS model and heterogeneous compound symmetry (CSH) model is that the CSH does not require the variances to be the same, which is exactly like the difference between AR(1) and ARH(1). The variance-covariance structure of CSH is expressed as following:

$$
\mathbf{V} = \begin{bmatrix}
\sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho & \cdots & \sigma_1\sigma_k\rho \\
\sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \cdots & \sigma_2\sigma_k\rho \\
\sigma_3\sigma_1\rho & \sigma_3\sigma_2\rho & \sigma_3^2 & \cdots & \vdots \\
\vdots & \vdots & \vdots & \ddots & \sigma_{k-2}\sigma_k\rho \\
\sigma_k\sigma_1\rho & \sigma_k\sigma_2\rho & \cdots & \sigma_k\sigma_{k-2}\rho & \sigma_k^2
\end{bmatrix}
$$

From the matrix form of CSH covariance structure above, we can see that the variances on the main diagonal of the matrix are different. And the CSH model also requires $k+1$ covariance parameters, where k represents the number of the observations per subject exactly like the ARH(1) model.

1.4.4 Mixed Model Analysis with UN Covariance Structure

Among all the covariance structures in the mixed model procedures, the unstructured (UN) covariance structure is the most complex one, because it places no requirement on the covariance structure. The UN structure allows both the variance terms and covariance terms to be different. The variance-covariance structure of UN is expressed as following:

$$
\mathbf{V} = \begin{bmatrix}
\sigma_1^2 & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1k} \\
\sigma_{12} & \sigma_2^2 & \sigma_{23} & \cdots & \sigma_{2k} \\
\sigma_{13} & \sigma_{23} & \sigma_3^2 & \cdots & \sigma_{3k} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\sigma_{1k} & \sigma_{2k} & \sigma_{3k} & \cdots & \sigma_k^2
\end{bmatrix}
$$

Since the UN structure places no conditions on the covariance structure, it requires fitting the most variance-covariance parameters of all the structures, that is $k(k+1)/2$.

1.5 Research Objective of My Report

In repeated measures design, if the time-dependent correlation is ignorable, the conventional approach univariate analysis of variance is often used to analyze the repeated measures data. This analysis basically treats the repeated measures data as the split-plot design, where the time is considered as the split-plot treatment. In this report, we examine treating the split-plot data, in which the assumption of compound symmetry does not hold, as a repeated measure design, and using the mixed models in the SAS system to describe the covariance structure of the split-plot data. We will also examine when the traditional split-plot analysis of a repeated measures design with an AR(1) variance-covariance structure is acceptable.

CHAPTER 2 - Real Data Example

2.1 Profile of the Example Data

Stelzleni *et al*. (2008) investigated the effects of concentrate feeding of cull beef cows and the muscle type on the tenderness of meat. Twenty-four cull beef cows were selected to conduct this experiment and were randomly assigned to a concentrate of feeding for 0, 42 or 84 days. We refer to this treatment factor as days on feed (DOF). From each beef cow, one steak from each of the nine muscle types (MT) was extracted. Each steak was cooked and several cores from each steak were taken. The Warner-Bratzler shear force (WBS) value, i.e. the force in kilograms required to shear the sample, was obtained on each core.

2.2 Examination of the Structure of the Experiment

The design of this experiment can be considered to be a split-plot with whole plots run in a completely randomized design. The whole plot treatment is DOF and its levels were randomly assigned to individual beef cows making a cow the whole plot experimental unit. For each whole plot experimental unit or cow, the levels of the subplot treatment factor, MT, are then observed by the selection of a steak from each of the nine muscle types. Here the subplot experimental unit is the steak and cores from each steak are subsamples. Strictly speaking, it is not possible to randomly assign the levels of MT individual steaks. Instead, the steaks are cut from the particular parts of the beef cow, and this makes them "random" samples of those particular muscle types. Because we cannot assign each of the nine steaks taken from an individual animal completely at random to a particular muscle type, one might argue that this experiment has no true randomization at the subplot treatment level, and the subplot measurements are more like repeated measures. However, this experiment does look a lot like a split-plot since it contains two levels of experimental units (the cow and the steak) with the smaller units always nested within the larger units. Thus, we could analyze it as a split-plot with completely randomized whole plots. Because some groups of muscle types are more similar to each other than types from another group, we suspect that the constant covariance term between subplot observations from compound symmetry may not hold here, especially because the subplot treatment factor contains a relatively large number of levels. Also, because some muscle types are tougher than

others, it would not be surprising if the assumption of a constant variance among subplot treatments were violated as well.

2.3 Examination of the Covariance Structure

The unstructured covariance structure places no restriction on the covariance structure. If the dataset is big enough, one could always try to fit the unstructured covariance structure first and have a brief idea about the covariance structure of the dataset. Table 2.1 is the covariance matrix gained by fitting the unstructured covariance structure in PROC MIXED to the example data. From this table we can see that, values on the diagonal, which are the variances, range from 0.60 to 3.17, and some of them are quite different from each other. This suggests possible heterogeneous covariance structure in the example data and the conventional split-plot approach which requires the compound symmetry covariance structure in the data may not be very proper here. Therefore, we consider using a repeated measures approach which incorporates this structure and the apparent relationships among the subplots to analyze the example data.

Table 2.1 The covariance matrix obtained by fitting the UN structure for the example data

MT	1	2	3	4	5	6	7	8	9
1	1.84	0.38	0.80	0.51	0.86	0.20	0.24	0.26	−0.002
2	0.38	0.60	0.11	0.04	0.06	0.07	0.35	0.06	−0.08
3	0.80	0.11	1.21	0.65	0.62	0.15	0.17	0.55	−0.06
4	0.51	0.04	0.65	3.17	1.11	0.36	0.42	0.83	−0.60
5	0.86	0.06	0.62	1.11	1.24	0.29	0.10	0.28	−0.15
6	0.20	0.07	0.15	0.36	0.29	0.23	−0.07	0.15	−0.18
7	0.24	0.35	0.17	0.42	0.10	−0.07	1.10	0.04	0.41
8	0.26	0.06	0.55	0.83	0.28	0.15	0.04	1.20	−0.31
9	−0.002	−0.08	−0.06	−0.60	−0.15	−0.18	0.41	−0.31	1.17

2.4 Different Covariance Structure Models in PROC MIXED

We fit the example data with five different covariance structure models in PROC MIXED (CS, CSH, AR(1), ARH(1), and UN). The CS covariance model yields the same AIC value as the conventional split-plot approach does. To use these five covariance structures in the SAS system, one just needs to add the 'TYPE =' option in the REPEATED statement in PROC

MIXED. For example, the SAS code for PROC MIXED with an AR(1) covariance structure is given below:

```
PROC MIXED DATA=COW;
    CLASS ANIMAL DOF MT;
    MODEL WBS=DOF MT DOF*MT;
    REPEATED MT/SUBJECT=ANIMAL(DOF) TYPE=AR(1) R RCORR;
RUN;
```

The AIC and BIC fit statistics for the five covariance models are shown in Table 2.2. From this table, we can see that both AIC and BIC values for the CSH covariance structure model are smaller than the corresponding AIC and BIC values for the rest covariance models. Therefore, we can conclude that the CSH covariance model fits the example data best among the five selected models and becomes our choice of model. Also note that UN and ARH(1), which also allow for unequal subplot treatment variances, both had smaller AIC values than CS.

Table 2.2 Model fit statistics with five different covariance structures for the example data

Fit statistics	Covariance Structures				
	CS	CSH	AR(1)	ARH(1)	UN
AIC	632.2	**609.3**	639.2	618.5	628.0
BIC	634.5	**621.1**	641.6	630.3	681.0

CHAPTER 3 - Simulation Study

In this section, the simulation study is conducted to evaluate the effectiveness of our proposed repeated measures approach in selecting the covariance structure of the simulated datasets. In this report, we consider two cases:

Case 1: We generate split-plot datasets with true CS covariance structure and then use CS and AR(1) covariance models in PROC MIXED, as well as the conventional split-plot model to analyze them and see how many times the conventional split-plot model can yield smaller AIC value than AR(1) covariance model.

Case 2: We generate split-plot datasets with AR(1) covariance structure and then use CS and AR(1) covariance models in PROC MIXED, as well as the conventional split-plot model to analyze them and see how many times the AR(1) covariance model can yield smaller AIC value than the conventional split-plot model.

For each case, we consider the number of whole plot treatment levels to be $a = 2$; number of split-plot treatment levels $b = 3$, 6, or 9; number of replications for whole plot units $r = 3$, 6, or 9; and variance ratio $\eta = \sigma_w^2 / \sigma_e^2 = 0.1$, 0.5, 1, 2, 10.

3.1 Simulation Study for Case 1

3.1.1 Simulation Procedure for Case1

For each selected setting:

 (1) Generate a split-plot in CRD dataset with CS covariance structure.

 (2) Use CS and AR(1) covariance models in PROC MIXED, as well as the conventional split-plot model to analyze the dataset been generated. For each model, we consider both default and Satterthwaite approximation method for computing the denominator degrees of freedom for the tests of fixed effects.

 (3) Repeat step (1) and (2) for $N = 2000$ times.

(4) Report the Type I error rate of the whole plot treatment main effect, subplot treatment main effect, and the interaction effect and also the percent of times that the AIC value of the conventional split-plot model is smaller than that of the AR(1) covariance model.

(5) Do step (1) to (4) for all the parameter settings.

(6) Display and summarize the results in the form of tables.

3.1.2 Simulation Results for Case 1

In the simulation study for case 1, we use R to generate the datasets we need and then export them to SAS to do the analysis using three different models. The simulation results are shown in Appendix A (Tables A.1 to A.9).

From Table A.1 to A.9, we can see that for each selected setting of a, b, r, n and N, as the value of η gets bigger, the percent of times that the AIC value of the conventional split-plot model is smaller than that of the AR(1) covariance model gets bigger. This suggests that when the whole plot error variance is significantly larger than the subplot error variance, the proposed repeated measures approach tends to have a larger chance to choose the true covariance structure of the data. From Table 3.1 below, we can see that as the sample size n increases, the chance that the proposed repeated measures approach has on selecting the true covariance structure of the data increases, and this is especially obvious when η is not very small. For the datasets with the same sample size, the proposed repeated measures approach works better on the dataset with a larger number of subplot treatment levels b.

As to the Type I error rate, from Table A.1 to A.9, we can see that for all the methods, the Type I error rates of the whole plot treatment main effect seem to be a little bit conservative. And as the value of η increases, it seems that the Type I error rates tend to get more conservative for most cases of the CS covariance model and quite a few cases of the conventional split-plot model. For the subplot treatment main effect and interaction effect, the Type I error rates of the CS covariance model and the conventional split-plot model are pretty good; while the Type I error rates of the AR(1) covariance model tend to be inflated a little bit, especially when η is big.

Table 3.1 Comparison of percent of times when AIC value of conventional split-plot method is smaller than that of AR(1) covariance model for different sample size for datasets with true CS covariance structure.

n	a	b	r	$\eta=0.1$	$\eta=0.5$	$\eta=1$	$\eta=2$	$\eta=10$
18	2	3	3	63.70%	60.05%	61.50%	67.10%	70.75%
36	2	3	6	63.90%	64.55%	72.85%	77.60%	80.80%
36	2	6	3	69.60%	74.50%	84.10%	90.15%	92.70%
54	2	3	9	67.90%	69.10%	78.70%	83.10%	86.15%
54	2	9	3	69.75%	80.85%	92.35%	97.40%	98.25%
72	2	6	6	67.30%	85.35%	96.00%	98.45%	98.70%
108	2	6	9	68.40%	91.70%	98.90%	99.70%	99.70%
108	2	9	6	67.00%	94.20%	99.55%	99.95%	99.95%
162	2	9	9	66.85%	98.05%	99.95%	100.00%	100.00%

3.2 Simulation Study for Case 2

3.2.1 Simulation Procedure for Case 2

For each selected setting:

     (1) Generate a split-plot in CRD dataset with AR(1) covariance structure.

     (2) Use CS and AR(1) covariance models in PROC MIXED, as well as the conventional split-plot model to analyze the dataset been generated. For each model, we consider both default and Satterthwaite approximation method for computing the denominator degrees of freedom for the tests of fixed effects.

     (3) Repeat step (1) and (2) for N=2000 times.

     (4) Report Type I error rate of the whole plot treatment main effect, subplot treatment main effect, and the interaction effect and also the percent of times that the AIC value of the AR(1) covariance model is smaller than that of the conventional split-plot model.

     (5) Do step (1) to (4) for all the parameter settings.

     (6) Display and summarize the results in the form of tables.

3.2.1 Simulation Result for Case 2

In the simulation study for case 2, we use SAS to generate the datasets we need through the method developed by Song, Xue and Li (2013) and then analyze them with three different models. The simulation results are shown in Appendix A (Tables A.10 to A.18).

From Table A.10 to A.18, we can see that for each selected setting of a, b, r, n and N, as the value of η gets bigger, the percent of times that the AIC value of the AR(1) covariance model is smaller than that of the conventional split-plot model gets bigger. This also suggests that when the whole plot error variance is significantly larger than the subplot error variance, the proposed repeated measures approach tends to have a larger chance to choose the true covariance structure of the data. From Table 3.2 below, we can see that as sample size n increases, the chance that the proposed repeated measures approach has on selecting the true covariance structure of the data increases, and this is especially obvious when η is not very small. For the datasets with the same sample size, the proposed repeated measures approach works better on the dataset with a larger number of b when η is not very small. However, unlike in Case 1, the proposed repeated measures approach consistently tends to choose the true covariance structure more than the untrue one, in Case 2, when η is very small, the proposed mixed model tends to choose the CS covariance structure instead of the true AR(1) covariance structure no matter how big the values of n and b are.

As to the Type I error rate, from Table A.10 to A.18, we can see that for all the methods, the Type I error rated of the whole plot treatment main effect seem to be acceptable, except that when η is very small, the conventional split-plot with Satterthwaith approximation method tends to be a little bit conservative than other methods. For the split-plot treatment main effect and interaction effect, when b is small, all methods have done similar work on controlling the Type I error rates; however as b gets bigger, the conventional split-plot model and the CS covariance model tend to inflate the Type I error rates, especially when η is very large.

Table 3.2 Comparison of percent of times when AIC value of AR(1) covariance model is smaller than that of conventional split-plot method for different sample size for datasets with true AR(1) covariance structure..

n	a	b	r	$\eta=0.1$	$\eta=0.5$	$\eta=1$	$\eta=2$	$\eta=10$
18	2	3	3	34.40%	47.35%	55.30%	63.80%	69.75%
36	2	3	6	37.40%	56.30%	68.85%	73.60%	79.55%
36	2	6	3	32.80%	59.50%	76.90%	83.75%	90.60%
54	2	3	9	38.25%	65.40%	74.55%	78.20%	85.45%
54	2	9	3	34.65%	70.05%	87.20%	93.75%	97.35%
72	2	6	6	39.10%	80.70%	91.50%	95.75%	98.25%
108	2	6	9	44.25%	86.75%	96.40%	98.10%	99.65%
108	2	9	6	42.95%	90.25%	97.75%	99.40%	99.90%
162	2	9	9	49.50%	96.65%	99.30%	99.95%	100.00%

CHAPTER 4 - Discussion

In this report, we investigated the use of mixed models in selecting the covariance structure of the split-plot data. Through the simulation study, we show that the proposed repeated measures approach works well on selecting the true covariance structure of the simulated datasets when the variance ratio η is not very small, and the sample size n and number of subplot treatment levels b is relatively big. And we also show that the true covariance model in both cases does a good job on controlling the Type I error rates of the subplot treatment main effect and interaction effect. In the real data example, we fit the data using five different covariance structure models in PROC MIXED (CS, CSH, AR(1), ARH(1), and UN) and it turns out that the CSH covariance model has the smallest values for the Akaike and Schwarz's Bayesian information criteria fit statistics and is therefore the model of our choice.

Even though in this report, we mainly focus on the split-plot in completely randomized design, the proposed repeated measures approach can also be applied to other split-plot designs, such as split-plot in randomized complete block design. We have applied the covariance structure models in PROC MIXED to other example data, and the results turn out to be very good.

There are some questions we need to take into consideration for future research. Firstly, in the simulation study, when the variance ratio η is very small, the simulation results are not that satisfying for both cases. One may consider other covariance models in PROC MIXED or analyzing the data as the completely randomized design since small η value may suggest that the whole plot error variance is so small that we can treat it as zero. Secondly, in the simulation study for Case 1, we find that for all the methods, the Type I error rates of the whole plot treatment seem to be a little bit conservative. Therefore it requires further study to properly explain and address those issues.

Bibliography

Box, G. E. P. and Jones S. (1990). Split-Plot Designs for Robust Product Experimentation. Journal of Applied Statistics, 19, 3-26.

Fisher, R. A. (1925). Statistical Methods for Research Workers. Edinburgh: Oliver and Boyd.

Jones, B. and Nachtsheim, C. J. (2009). Split-Plot Designs: What, Why, and How. Journal of Quality Technology, 41, 340-361.

Kowalski, S. M. and Potcner, K. J. (2003). How to recognize a split-plot experiment. Quality Progress, 36(11), 60-66.

Littell, R. C., Milliken, G. A., and Stroup, W. W., and Wolfinger, R. D. (1996). The SAS System for Mixed Models. Cary, NC: SAS Institute.

Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., and Schabenberger, O. (2006). SAS for Mixed Models. Cary, NC: SAS Institute.

Littell, R. C., Pendergast, J., and Natarajan, R. (2000). Tutorial in biostatistics–modelling covariance structure in the analysis of repeated measures data. *Statistics in medicine*, *19*(1793), 1819.

Mbegbu, J. and Francis, O. I. (2012). Comparison of the Optimal Design: Split-Plot Experiments. Mathematical Theory and Modeling, 2(7), 23-31.

Song, P., Xue, J., and Li, Z. (2013). Simulation of Longitudinal Exposure Data with Variance-Covariance Structures Based on Mixed Models. *Risk Analysis*, *33*(3), 469-479.

Stelzleni, A. M., Johnson, D. D., and Thrift, T. A. (2008). Effects of Days on Concentrate Feed and Postmortem Aging on Carcass and Palatability Characteristics of Selected Muscles from Cull Beef Cows. The Professional Animal Scientist, 24, 334-341.

Vahl, C. I. and Milliken, G. A. (2011). Whole-plot Exchange Algorithms for Constructing D-Optimal Multistratum Designs. Taylor & Francis, 40, 1030-1042.

Wolfinger, R., Tobias, R., and Sall, J. (1991). Mixed Models: A Future Direction. SUGI, 16, 1380-1388.

Wang, Z. and Goonewardene, L. A. (2004). The use of MIXED models in the analysis of animal experiments with repeated measures data. Canadian Journal of Animal Science, 84(1), 1-11.

Appendix A - Simulation Results for Case 1 and 2

Table A.1 Comparison of Type I error rate and AIC value of CS and AR(1) covariance mixed model methods, and conventional split-plot method for datasets with true CS covariance structure and where $a=2$, $b=3$, $r=3$, $n=18$, and $N=2000$.

Method	[3]Source	[1]$a=2$; $b=3$; $r=3$; $n=18$; $N=2000$				
		$\eta=0.1$	$\eta=0.5$	$\eta=1$	$\eta=2$	$\eta=10$
Split-plot	WP	0.033	0.027	0.033	0.021	0.023
	SP	0.058	0.061	0.045	0.052	0.057
	WP*SP	0.054	0.060	0.055	0.043	0.056
Split-plot/satterth	WP	0.016	0.020	0.028	0.021	0.023
	SP	0.054	0.059	0.045	0.052	0.057
	WP*SP	0.048	0.055	0.055	0.043	0.056
CS	WP	0.050	0.031	0.033	0.021	0.023
	SP	0.047	0.058	0.045	0.052	0.057
	WP*SP	0.044	0.053	0.055	0.043	0.056
CS/satterth	WP	0.050	0.031	0.033	0.021	0.023
	SP	0.047	0.058	0.045	0.052	0.057
	WP*SP	0.044	0.053	0.055	0.043	0.056
AR(1)	WP	0.045	0.039	0.037	0.024	0.023
	SP	0.056	0.065	0.049	0.060	0.069
	WP*SP	0.058	0.051	0.059	0.052	0.068
AR(1)/satterth	WP	0.029	0.026	0.033	0.023	0.023
	SP	0.057	0.065	0.051	0.058	0.069
	WP*SP	0.056	0.054	0.059	0.053	0.068
Percent of times AIC of [2]SP<AIC of AR(1)		63.70%	60.05%	61.50%	67.10%	70.75%

[1]a is the whole plot treatment level, b is the split-plot treatment level, r is number of replications for whole plot units, n is the sample size, and N is repetitions.

[2]SP is the conventional split-plot method. AIC of SP is exactly the same as AIC of CS model.

[3]Source: WP is whole plot main effect, SP is subplot main effect, WP*SP is interaction effect.

Table A.2 Comparison of Type I error rate and AIC value of CS and AR(1) covariance mixed model methods, and conventional split-plot method for datasets with true CS covariance structure and where $a=2$, $b=6$, $r=3$, $n=36$, and $N=2000$.

Method	[3]Source	[1]$a=2$; $b=6$; $r=3$; $n=36$; $N=2000$				
		$\eta=0.1$	$\eta=0.5$	$\eta=1$	$\eta=2$	$\eta=10$
Split-plot	WP	0.026	0.031	0.026	0.017	0.022
	SP	0.054	0.056	0.058	0.046	0.046
	WP*SP	0.058	0.051	0.054	0.047	0.058
Split-plot/satterth	WP	0.008	0.018	0.025	0.017	0.022
	SP	0.052	0.055	0.058	0.046	0.046
	WP*SP	0.055	0.051	0.054	0.047	0.058
CS	WP	0.048	0.028	0.026	0.017	0.022
	SP	0.046	0.053	0.058	0.046	0.046
	WP*SP	0.051	0.050	0.053	0.046	0.058
CS/satterth	WP	0.048	0.028	0.026	0.017	0.022
	SP	0.046	0.053	0.058	0.046	0.046
	WP*SP	0.051	0.050	0.053	0.046	0.058
AR(1)	WP	0.039	0.065	0.060	0.037	0.024
	SP	0.053	0.052	0.061	0.065	0.071
	WP*SP	0.047	0.044	0.063	0.072	0.081
AR(1)/satterth	WP	0.016	0.032	0.045	0.031	0.023
	SP	0.061	0.061	0.064	0.066	0.071
	WP*SP	0.059	0.050	0.065	0.073	0.081
Percent of times AIC of [2]SP <AIC of AR(1)		69.60%	74.50%	84.10%	90.15%	92.70%

[1]a is the whole plot treatment level, b is the split-plot treatment level, r is number of replications for whole plot units, n is the sample size, and N is repetitions.

[2]SP is the conventional split-plot method. AIC of SP is exactly the same as AIC of CS model.

[3]Source: WP is whole plot main effect, SP is subplot main effect, WP*SP is interaction effect.

Table A.3 Comparison of Type I error rate and AIC value of CS and AR(1) covariance mixed model methods, and conventional split-plot method for datasets with true CS covariance structure and where *a*=2, *b*=9, *r*=3, *n*=54, and *N*=2000.

Method	[3]Source	[1]a=2; b=9; r=3; n=54; N=2000				
		η=0.1	η=0.5	η=1	η=2	η=10
Split-plot	WP	0.029	0.029	0.022	0.017	0.023
	SP	0.066	0.061	0.048	0.057	0.054
	WP*SP	0.055	0.058	0.053	0.054	0.053
Split-plot/satterth	WP	0.007	0.018	0.022	0.017	0.023
	SP	0.065	0.061	0.048	0.057	0.054
	WP*SP	0.054	0.058	0.053	0.054	0.053
CS	WP	0.042	0.027	0.022	0.017	0.023
	SP	0.060	0.061	0.048	0.057	0.054
	WP*SP	0.050	0.057	0.053	0.054	0.053
CS/satterth	WP	0.042	0.027	0.022	0.017	0.023
	SP	0.060	0.061	0.048	0.057	0.054
	WP*SP	0.050	0.057	0.053	0.054	0.053
AR(1)	WP	0.041	0.076	0.081	0.054	0.027
	SP	0.049	0.046	0.049	0.060	0.074
	WP*SP	0.045	0.044	0.044	0.066	0.086
AR(1)/satterth	WP	0.014	0.037	0.055	0.041	0.026
	SP	0.072	0.052	0.054	0.065	0.074
	WP*SP	0.062	0.054	0.049	0.068	0.086
Percent of times AIC of [2]SP <AIC of AR(1)		69.75%	80.85%	92.35%	97.40%	98.25%

[1]a is the whole plot treatment level, b is the split-plot treatment level, r is number of replications for whole plot units, n is the sample size, and N is repetitions.

[2]SP is the conventional split-plot method. AIC of SP is exactly the same as AIC of CS model.

[3]Source: WP is whole plot main effect, SP is subplot main effect, WP*SP is interaction effect.

Table A.4 Comparison of Type I error rate and AIC value of CS and AR(1) covariance mixed model methods, and conventional split-plot method for datasets with true CS covariance structure and where $a=2$, $b=3$, $r=6$, $n=36$, and $N=2000$.

Method	[3]Source	[1]$a=2$; $b=3$; $r=6$; $n=36$; $N=2000$				
		$\eta=0.1$	$\eta=0.5$	$\eta=1$	$\eta=2$	$\eta=10$
Split-plot	WP	0.029	0.021	0.020	0.015	0.018
	SP	0.050	0.047	0.055	0.052	0.049
	WP*SP	0.055	0.049	0.057	0.049	0.047
Split-plot/satterth	WP	0.023	0.020	0.020	0.015	0.018
	SP	0.046	0.047	0.055	0.052	0.049
	WP*SP	0.048	0.048	0.057	0.049	0.047
CS	WP	0.042	0.023	0.020	0.015	0.018
	SP	0.043	0.047	0.055	0.052	0.049
	WP*SP	0.044	0.048	0.057	0.049	0.047
CS/satterth	WP	0.042	0.023	0.020	0.015	0.018
	SP	0.043	0.047	0.055	0.052	0.049
	WP*SP	0.044	0.048	0.057	0.049	0.047
AR(1)	WP	0.038	0.031	0.024	0.015	0.018
	SP	0.048	0.051	0.047	0.065	0.063
	WP*SP	0.049	0.052	0.054	0.052	0.058
AR(1)/satterth	WP	0.032	0.029	0.024	0.015	0.018
	SP	0.049	0.050	0.049	0.065	0.063
	WP*SP	0.048	0.051	0.054	0.052	0.058
Percent of times AIC of [2]SP <AIC of AR(1)		63.90%	64.55%	72.85%	77.60%	80.80%

[1]a is the whole plot treatment level, b is the split-plot treatment level, r is number of replications for whole plot units, n is the sample size, and N is repetitions.

[2]SP is the conventional split-plot method. AIC of SP is exactly the same as AIC of CS model.

[3]Source: WP is whole plot main effect, SP is subplot main effect, WP*SP is interaction effect.

Table A.5 Comparison of Type I error rate and AIC value of CS and AR(1) covariance mixed model methods, and conventional split-plot method for datasets with true CS covariance structure and where *a*=2, *b*=6, *r*=6, *n*=72, and *N*=2000.

Method	[3]Source	[1]*a*=2; *b*=6; *r*=6; *n*=72; *N*=2000				
		η=0.1	η=0.5	η=1	η=2	η=10
	WP	0.030	0.015	0.012	0.014	0.015
Split-plot	SP	0.056	0.048	0.049	0.054	0.054
	WP*SP	0.053	0.052	0.062	0.044	0.057
	WP	0.023	0.015	0.012	0.014	0.015
Split-plot/satterth	SP	0.056	0.048	0.049	0.054	0.054
	WP*SP	0.053	0.052	0.062	0.044	0.057
	WP	0.045	0.015	0.012	0.014	0.015
CS	SP	0.055	0.048	0.049	0.054	0.054
	WP*SP	0.050	0.052	0.062	0.044	0.057
	WP	0.045	0.015	0.012	0.014	0.015
CS/satterth	SP	0.055	0.048	0.049	0.054	0.054
	WP*SP	0.050	0.052	0.062	0.044	0.057
	WP	0.049	0.054	0.045	0.026	0.016
AR(1)	SP	0.051	0.041	0.058	0.061	0.076
	WP*SP	0.043	0.044	0.062	0.057	0.085
	WP	0.036	0.045	0.039	0.024	0.016
AR(1)/satterth	SP	0.055	0.042	0.060	0.062	0.076
	WP*SP	0.050	0.046	0.064	0.057	0.085
Percent of times AIC of [2]SP <AIC of AR(1)		67.30%	85.35%	96.00%	98.45%	98.70%

[1]*a* is the whole plot treatment level, *b* is the split-plot treatment level, *r* is number of replications for whole plot units, *n* is the sample size, and *N* is repetitions.

[2]SP is the conventional split-plot method. AIC of SP is exactly the same as AIC of CS model.

[3]Source: WP is whole plot main effect, SP is subplot main effect, WP*SP is interaction effect.

Table A.6 Comparison of Type I error rate and AIC value of CS and AR(1) covariance mixed model methods, and conventional split-plot method for datasets with true CS covariance structure and where $a=2$, $b=9$, $r=6$, $n=108$, and $N=2000$.

Method	[3]Source	[1]$a=2$; $b=9$; $r=6$; $n=108$; $N=2000$				
		$\eta=0.1$	$\eta=0.5$	$\eta=1$	$\eta=2$	$\eta=10$
Split-plot	WP	0.029	0.018	0.017	0.012	0.011
	SP	0.057	0.046	0.065	0.053	0.046
	WP*SP	0.055	0.044	0.047	0.048	0.050
Split-plot/satterth	WP	0.018	0.017	0.017	0.012	0.011
	SP	0.057	0.046	0.065	0.053	0.046
	WP*SP	0.054	0.044	0.047	0.048	0.050
CS	WP	0.037	0.018	0.017	0.012	0.011
	SP	0.053	0.046	0.065	0.053	0.046
	WP*SP	0.053	0.044	0.047	0.048	0.050
CS/satterth	WP	0.037	0.018	0.017	0.012	0.011
	SP	0.053	0.046	0.065	0.053	0.046
	WP*SP	0.053	0.044	0.047	0.048	0.050
AR(1)	WP	0.042	0.080	0.058	0.028	0.012
	SP	0.048	0.035	0.052	0.063	0.069
	WP*SP	0.046	0.030	0.047	0.060	0.073
AR(1)/satterth	WP	0.027	0.059	0.050	0.025	0.012
	SP	0.056	0.039	0.055	0.065	0.069
	WP*SP	0.050	0.033	0.050	0.061	0.074
Percent of times AIC of [2]SP <AIC of AR(1)		67.00%	94.20%	99.55%	99.95%	99.95%

[1]a is the whole plot treatment level, b is the split-plot treatment level, r is number of replications for whole plot units, n is the sample size, and N is repetitions.

[2]SP is the conventional split-plot method. AIC of SP is exactly the same as AIC of CS model.

[3]Source: WP is whole plot main effect, SP is subplot main effect, WP*SP is interaction effect.

Table A.7 Comparison of Type I error rate and AIC value of CS and AR(1) covariance mixed model methods, and conventional split-plot method for datasets with true CS covariance structure and where $a=2$, $b=3$, $r=9$, $n=54$, and $N=2000$.

Method	[3]Source	[1]$a=2$; $b=3$; $r=9$; $n=54$; $N=2000$				
		$\eta=0.1$	$\eta=0.5$	$\eta=1$	$\eta=2$	$\eta=10$
	WP	0.038	0.032	0.015	0.009	0.007
Split-plot	SP	0.050	0.053	0.055	0.045	0.060
	WP*SP	0.061	0.045	0.050	0.053	0.051
	WP	0.032	0.032	0.015	0.009	0.007
Split-plot/satterth	SP	0.050	0.053	0.055	0.045	0.060
	WP*SP	0.059	0.045	0.050	0.053	0.051
	WP	0.050	0.032	0.015	0.009	0.007
CS	SP	0.044	0.053	0.055	0.045	0.060
	WP*SP	0.052	0.044	0.050	0.053	0.051
	WP	0.050	0.032	0.015	0.009	0.007
CS/satterth	SP	0.044	0.053	0.055	0.045	0.060
	WP*SP	0.052	0.044	0.050	0.053	0.051
	WP	0.049	0.039	0.021	0.010	0.007
AR(1)	SP	0.048	0.049	0.055	0.048	0.063
	WP*SP	0.055	0.043	0.045	0.060	0.059
	WP	0.042	0.038	0.020	0.009	0.007
AR(1)/satterth	SP	0.048	0.049	0.055	0.048	0.063
	WP*SP	0.054	0.042	0.045	0.061	0.059
Percent of times AIC of [2]SP <AIC of AR(1)		67.90%	69.10%	78.70%	83.10%	86.15%

[1]a is the whole plot treatment level, b is the split-plot treatment level, r is number of replications for whole plot units, n is the sample size, and N is repetitions.

[2]SP is the conventional split-plot method. AIC of SP is exactly the same as AIC of CS model.

[3]Source: WP is whole plot main effect, SP is subplot main effect, WP*SP is interaction effect.

Table A.8 Comparison of Type I error rate and AIC value of CS and AR(1) covariance mixed model methods, and conventional split-plot method for datasets with true CS covariance structure and where $a=2$, $b=6$, $r=9$, $n=108$, and $N=2000$.

Method	[3]Source	[1]$a=2$; $b=6$; $r=9$; $n=108$; $N=2000$				
		$\eta=0.1$	$\eta=0.5$	$\eta=1$	$\eta=2$	$\eta=10$
Split-plot	WP	0.034	0.018	0.013	0.008	0.010
	SP	0.059	0.056	0.047	0.043	0.050
	WP*SP	0.050	0.054	0.057	0.052	0.051
Split-plot/satterth	WP	0.028	0.018	0.013	0.008	0.010
	SP	0.058	0.056	0.047	0.043	0.050
	WP*SP	0.050	0.054	0.057	0.052	0.051
CS	WP	0.039	0.018	0.013	0.008	0.010
	SP	0.054	0.056	0.047	0.043	0.050
	WP*SP	0.048	0.054	0.057	0.052	0.051
CS/satterth	WP	0.039	0.018	0.013	0.008	0.010
	SP	0.054	0.056	0.047	0.043	0.050
	WP*SP	0.048	0.054	0.057	0.052	0.051
AR(1)	WP	0.040	0.050	0.036	0.015	0.011
	SP	0.053	0.039	0.046	0.056	0.075
	WP*SP	0.049	0.040	0.049	0.060	0.067
AR(1)/satterth	WP	0.031	0.042	0.033	0.014	0.011
	SP	0.058	0.041	0.047	0.056	0.075
	WP*SP	0.051	0.042	0.049	0.060	0.067
Percent of times AIC of [2]SP <AIC of AR(1)		68.40%	91.70%	98.90%	99.70%	99.70%

[1]a is the whole plot treatment level, b is the split-plot treatment level, r is number of replications for whole plot units, n is the sample size, and N is repetitions.

[2]SP is the conventional split-plot method. AIC of SP is exactly the same as AIC of CS model.

[3]Source: WP is whole plot main effect, SP is subplot main effect, WP*SP is interaction effect.

Table A.9 Comparison of Type I error rate and AIC value of CS and AR(1) covariance mixed model methods, and conventional split-plot method for datasets with true CS covariance structure and where $a=2$, $b=9$, $r=9$, $n=162$, and $N=2000$.

Method	[3]Source	[1]$a=2$; $b=9$; $r=9$; $n=162$; $N=2000$				
		$\eta=0.1$	$\eta=0.5$	$\eta=1$	$\eta=2$	$\eta=10$
	WP	0.033	0.022	0.014	0.014	0.010
Split-plot	SP	0.052	0.051	0.052	0.058	0.048
	WP*SP	0.050	0.053	0.058	0.051	0.049
	WP	0.028	0.022	0.014	0.014	0.010
Split-plot/satterth	SP	0.051	0.051	0.052	0.058	0.048
	WP*SP	0.050	0.054	0.058	0.051	0.049
	WP	0.041	0.022	0.014	0.014	0.010
CS	SP	0.049	0.051	0.052	0.058	0.048
	WP*SP	0.049	0.054	0.058	0.051	0.049
	WP	0.041	0.022	0.014	0.014	0.010
CS/satterth	SP	0.049	0.051	0.052	0.058	0.048
	WP*SP	0.049	0.054	0.058	0.051	0.049
	WP	0.055	0.082	0.058	0.026	0.010
AR(1)	SP	0.044	0.032	0.044	0.067	0.074
	WP*SP	0.046	0.034	0.050	0.070	0.081
	WP	0.045	0.070	0.052	0.025	0.010
AR(1)/satterth	SP	0.048	0.034	0.047	0.067	0.074
	WP*SP	0.049	0.035	0.052	0.070	0.081
Percent of times AIC of [2]SP <AIC of AR(1)		66.85%	98.05%	99.95%	100.00%	100.00%

[1]a is the whole plot treatment level, b is the split-plot treatment level, r is number of replications for whole plot units, n is the sample size, and N is repetitions.

[2]SP is the conventional split-plot method. AIC of SP is exactly the same as AIC of CS model.

[3]Source: WP is whole plot main effect, SP is subplot main effect, WP*SP is interaction effect.

Table A.10 Comparison of Type I error rate and AIC value of CS and AR(1) covariance mixed model methods, and conventional split-plot method for datasets with true AR(1) covariance structure and where $a=2$, $b=3$, $r=3$, $n=18$, and $N=2000$.

Method	[3]Source	[1]$a=2$; $b=3$; $r=3$; $n=18$; $N=2000$				
		$\eta=0.1$	$\eta=0.5$	$\eta=1$	$\eta=2$	$\eta=10$
	WP	0.032	0.042	0.056	0.055	0.048
Split-plot	SP	0.068	0.057	0.053	0.052	0.061
	WP*SP	0.064	0.053	0.061	0.063	0.055
	WP	0.014	0.028	0.046	0.046	0.047
Split-plot/satterth	SP	0.062	0.053	0.051	0.052	0.061
	WP*SP	0.057	0.051	0.059	0.062	0.055
	WP	0.043	0.050	0.058	0.054	0.047
CS	SP	0.057	0.051	0.050	0.052	0.061
	WP*SP	0.051	0.051	0.057	0.062	0.055
	WP	0.043	0.050	0.058	0.054	0.047
CS/satterth	SP	0.057	0.051	0.050	0.052	0.061
	WP*SP	0.051	0.051	0.057	0.062	0.055
	WP	0.037	0.041	0.057	0.052	0.047
AR(1)	SP	0.062	0.057	0.052	0.045	0.056
	WP*SP	0.053	0.046	0.050	0.056	0.052
	WP	0.024	0.032	0.050	0.046	0.047
AR(1)/satterth	SP	0.060	0.058	0.053	0.048	0.056
	WP*SP	0.054	0.048	0.053	0.057	0.052
Percent of times AIC of AR(1)<AIC of [2]SP		34.40%	47.35%	55.30%	63.80%	69.75%

[1]a is the whole plot treatment level, b is the split-plot treatment level, r is number of replications for whole plot units, n is the sample size, and N is repetitions.

[2]SP is the conventional split-plot method. AIC of SP is exactly the same as AIC of CS model.

[3]Source: WP is whole plot main effect, SP is subplot main effect, WP*SP is interaction effect.

Table A.11 Comparison of Type I error rate and AIC value of CS and AR(1) covariance mixed model methods, and conventional split-plot method for datasets with true AR(1) covariance structure and where $a=2$, $b=6$, $r=3$, $n=18$, and $N=2000$.

| Method | [3]Source | [1]$a=2$; $b=6$; $r=3$; $n=36$; $N=2000$ | | | | |
		$\eta=0.1$	$\eta=0.5$	$\eta=1$	$\eta=2$	$\eta=10$
	WP	0.037	0.048	0.052	0.040	0.057
Split-plot	SP	0.058	0.055	0.058	0.085	0.094
	WP*SP	0.058	0.067	0.059	0.099	0.098
	WP	0.010	0.022	0.034	0.035	0.056
Split-plot/satterth	SP	0.055	0.054	0.058	0.085	0.094
	WP*SP	0.055	0.067	0.058	0.098	0.098
	WP	0.050	0.054	0.050	0.041	0.057
CS	SP	0.052	0.052	0.057	0.084	0.094
	WP*SP	0.049	0.064	0.057	0.098	0.098
	WP	0.050	0.054	0.050	0.041	0.057
CS/satterth	SP	0.052	0.052	0.057	0.084	0.094
	WP*SP	0.049	0.064	0.057	0.098	0.098
	WP	0.044	0.043	0.043	0.039	0.060
AR(1)	SP	0.051	0.051	0.040	0.057	0.045
	WP*SP	0.050	0.060	0.049	0.068	0.059
	WP	0.019	0.019	0.024	0.027	0.048
AR(1)/satterth	SP	0.062	0.059	0.048	0.062	0.046
	WP*SP	0.060	0.067	0.058	0.073	0.060
Percent of times AIC of AR(1)<AIC of [2]SP		32.80%	59.50%	76.90%	83.75%	90.60%

[1]a is the whole plot treatment level, b is the split-plot treatment level, r is number of replications for whole plot units, n is the sample size, and N is repetitions.

[2]SP is the conventional split-plot method. AIC of SP is exactly the same as AIC of CS model.

[3]Source: WP is whole plot main effect, SP is subplot main effect, WP*SP is interaction effect.

Table A.12 Comparison of Type I error rate and AIC value of CS and AR(1) covariance mixed model methods, and conventional split-plot method for datasets with true AR(1) covariance structure and where $a=2$, $b=9$, $r=3$, $n=54$, and $N=2000$.

Method	[3]Source	[1]$a=2$; $b=9$; $r=3$; $n=54$; $N=2000$				
		$\eta=0.1$	$\eta=0.5$	$\eta=1$	$\eta=2$	$\eta=10$
Split-plot	WP	0.038	0.051	0.051	0.052	0.050
	SP	0.060	0.060	0.075	0.100	0.125
	WP*SP	0.062	0.066	0.067	0.102	0.130
Split-plot/satterth	WP	0.009	0.023	0.036	0.041	0.050
	SP	0.057	0.060	0.075	0.100	0.125
	WP*SP	0.060	0.065	0.067	0.102	0.130
CS	WP	0.051	0.048	0.047	0.050	0.050
	SP	0.051	0.059	0.074	0.100	0.125
	WP*SP	0.054	0.064	0.065	0.102	0.130
CS/satterth	WP	0.051	0.048	0.047	0.050	0.050
	SP	0.051	0.060	0.074	0.100	0.125
	WP*SP	0.054	0.064	0.065	0.102	0.130
AR(1)	WP	0.044	0.044	0.043	0.046	0.044
	SP	0.044	0.047	0.047	0.049	0.058
	WP*SP	0.049	0.055	0.043	0.051	0.054
AR(1)/satterth	WP	0.015	0.017	0.019	0.023	0.035
	SP	0.061	0.061	0.058	0.054	0.059
	WP*SP	0.062	0.066	0.052	0.062	0.055
Percent of times AIC of AR(1)<AIC of [2]SP		34.65%	70.05%	87.20%	93.75%	97.35%

[1]a is the whole plot treatment level, b is the split-plot treatment level, r is number of replications for whole plot units, n is the sample size, and N is repetitions.

[2]SP is the conventional split-plot method. AIC of SP is exactly the same as AIC of CS model.

[3]Source: WP is whole plot main effect, SP is subplot main effect, WP*SP is interaction effect.

Table A.13 Comparison of Type I error rate and AIC value of CS and AR(1) covariance mixed model methods, and conventional split-plot method for datasets with true AR(1) covariance structure and where $a=2$, $b=3$, $r=6$, $n=36$, and $N=2000$.

Method	[3]Source	[1]$a=2$; $b=3$; $r=6$; $n=36$; $N=2000$				
		$\eta=0.1$	$\eta=0.5$	$\eta=1$	$\eta=2$	$\eta=10$
	WP	0.037	0.046	0.043	0.058	0.053
Split-plot	SP	0.053	0.055	0.050	0.059	0.058
	WP*SP	0.060	0.050	0.065	0.054	0.056
	WP	0.031	0.043	0.043	0.058	0.053
Split-plot/satterth	SP	0.050	0.054	0.050	0.059	0.058
	WP*SP	0.057	0.049	0.065	0.054	0.056
	WP	0.048	0.048	0.044	0.058	0.053
CS	SP	0.046	0.054	0.049	0.059	0.058
	WP*SP	0.053	0.048	0.065	0.054	0.056
	WP	0.048	0.048	0.044	0.058	0.053
CS/satterth	SP	0.046	0.054	0.049	0.059	0.058
	WP*SP	0.053	0.048	0.065	0.054	0.056
	WP	0.043	0.047	0.042	0.057	0.054
AR(1)	SP	0.051	0.057	0.054	0.057	0.055
	WP*SP	0.059	0.049	0.060	0.049	0.049
	WP	0.036	0.041	0.039	0.054	0.054
AR(1)/satterth	SP	0.051	0.057	0.053	0.059	0.055
	WP*SP	0.058	0.048	0.060	0.050	0.049
Percent of times AIC of AR(1)<AIC of [2]SP		37.40%	56.30%	68.85%	73.60%	79.55%

[1]a is the whole plot treatment level, b is the split-plot treatment level, r is number of replications for whole plot units, n is the sample size, and N is repetitions.

[2]SP is the conventional split-plot method. AIC of SP is exactly the same as AIC of CS model.

[3]Source: WP is whole plot main effect, SP is subplot main effect, WP*SP is interaction effect.

Table A.14 Comparison of Type I error rate and AIC value of CS and AR(1) covariance mixed model methods, and conventional split-plot method for datasets with true AR(1) covariance structure and where $a=2$, $b=6$, $r=6$, $n=72$, and $N=2000$.

Method	[3]Source	[1]$a=2$; $b=6$; $r=6$; $n=72$; $N=2000$				
		$\eta=0.1$	$\eta=0.5$	$\eta=1$	$\eta=2$	$\eta=10$
Split-plot	WP	0.034	0.051	0.063	0.046	0.050
	SP	0.055	0.059	0.070	0.069	0.100
	WP*SP	0.047	0.060	0.060	0.069	0.095
Split-plot/satterth	WP	0.025	0.049	0.062	0.046	0.050
	SP	0.054	0.059	0.070	0.069	0.100
	WP*SP	0.047	0.059	0.060	0.069	0.095
CS	WP	0.044	0.055	0.063	0.046	0.050
	SP	0.048	0.058	0.070	0.069	0.100
	WP*SP	0.045	0.059	0.060	0.069	0.095
CS/satterth	WP	0.044	0.055	0.063	0.046	0.050
	SP	0.048	0.058	0.070	0.069	0.100
	WP*SP	0.045	0.059	0.060	0.069	0.095
AR(1)	WP	0.045	0.049	0.060	0.047	0.050
	SP	0.048	0.057	0.050	0.051	0.044
	WP*SP	0.042	0.054	0.043	0.047	0.052
AR(1)/satterth	WP	0.029	0.042	0.050	0.041	0.049
	SP	0.058	0.064	0.053	0.052	0.044
	WP*SP	0.046	0.057	0.047	0.049	0.052
Percent of times AIC of AR(1)<AIC of [2]SP		39.10%	80.70%	91.50%	95.75%	98.25%

[1]a is the whole plot treatment level, b is the split-plot treatment level, r is number of replications for whole plot units, n is the sample size, and N is repetitions.

[2]SP is the conventional split-plot method. AIC of SP is exactly the same as AIC of CS model.

[3]Source: WP is whole plot main effect, SP is subplot main effect, WP*SP is interaction effect.

Table A.15 Comparison of Type I error rate and AIC value of CS and AR(1) covariance mixed model methods, and conventional split-plot method for datasets with true AR(1) covariance structure and where $a=2$, $b=9$, $r=6$, $n=108$, and $N=2000$.

Method	[3]Source	[1]$a=2$; $b=9$; $r=6$; $n=108$; $N=2000$				
		$\eta=0.1$	$\eta=0.5$	$\eta=1$	$\eta=2$	$\eta=10$
Split-plot	WP	0.037	0.051	0.055	0.048	0.052
	SP	0.051	0.067	0.079	0.091	0.118
	WP*SP	0.055	0.062	0.071	0.092	0.120
Split-plot/satterth	WP	0.026	0.044	0.054	0.048	0.052
	SP	0.050	0.067	0.079	0.091	0.118
	WP*SP	0.055	0.061	0.071	0.092	0.120
CS	WP	0.049	0.050	0.055	0.048	0.052
	SP	0.047	0.066	0.079	0.091	0.118
	WP*SP	0.051	0.060	0.071	0.092	0.120
CS/satterth	WP	0.049	0.050	0.055	0.048	0.052
	SP	0.047	0.066	0.079	0.091	0.118
	WP*SP	0.051	0.060	0.071	0.092	0.120
AR(1)	WP	0.049	0.048	0.053	0.045	0.047
	SP	0.047	0.049	0.054	0.052	0.048
	WP*SP	0.048	0.046	0.039	0.051	0.047
AR(1)/satterth	WP	0.031	0.035	0.040	0.037	0.045
	SP	0.051	0.057	0.059	0.054	0.050
	WP*SP	0.053	0.050	0.045	0.055	0.048
Percent of times AIC of AR(1)<AIC of [2]SP		42.95%	90.25%	97.75%	99.40%	99.90%

[1]a is the whole plot treatment level, b is the split-plot treatment level, r is number of replications for whole plot units, n is the sample size, and N is repetitions.

[2]SP is the conventional split-plot method. AIC of SP is exactly the same as AIC of CS model.

[3]Source: WP is whole plot main effect, SP is subplot main effect, WP*SP is interaction effect.

Table A.16 Comparison of Type I error rate and AIC value of CS and AR(1) covariance mixed model methods, and conventional split-plot method for datasets with true AR(1) covariance structure and where $a=2$, $b=3$, $r=9$, $n=54$, and $N=2000$.

Method	[3]Source	[1]$a=2$; $b=3$; $r=9$; $n=54$; $N=2000$				
		$\eta=0.1$	$\eta=0.5$	$\eta=1$	$\eta=2$	$\eta=10$
Split-plot	WP	0.041	0.051	0.053	0.052	0.046
	SP	0.056	0.045	0.041	0.058	0.061
	WP*SP	0.052	0.055	0.051	0.051	0.058
Split-plot/satterth	WP	0.034	0.050	0.053	0.052	0.046
	SP	0.053	0.045	0.041	0.058	0.061
	WP*SP	0.049	0.055	0.051	0.051	0.058
CS	WP	0.049	0.051	0.053	0.052	0.046
	SP	0.051	0.044	0.041	0.058	0.061
	WP*SP	0.047	0.055	0.051	0.051	0.058
CS/satterth	WP	0.049	0.051	0.053	0.052	0.046
	SP	0.051	0.044	0.041	0.058	0.061
	WP*SP	0.047	0.055	0.051	0.051	0.058
AR(1)	WP	0.040	0.048	0.055	0.052	0.048
	SP	0.055	0.044	0.044	0.056	0.056
	WP*SP	0.056	0.059	0.046	0.047	0.050
AR(1)/satterth	WP	0.035	0.047	0.054	0.051	0.047
	SP	0.055	0.044	0.044	0.055	0.056
	WP*SP	0.056	0.059	0.046	0.047	0.050
Percent of times AIC of AR(1)<AIC of [2]SP		38.25%	65.40%	74.55%	78.20%	85.45%

[1]a is the whole plot treatment level, b is the split-plot treatment level, r is number of replications for whole plot units, n is the sample size, and N is repetitions.

[2]SP is the conventional split-plot method. AIC of SP is exactly the same as AIC of CS model.

[3]Source: WP is whole plot main effect, SP is subplot main effect, WP*SP is interaction effect.

Table A.17 Comparison of Type I error rate and AIC value of CS and AR(1) covariance mixed model methods, and conventional split-plot method for datasets with true AR(1) covariance structure and where $a=2$, $b=6$, $r=9$, $n=108$, and $N=2000$.

Method	[3]Source	[1]$a=2$; $b=6$; $r=9$; $n=108$; $N=2000$				
		$\eta=0.1$	$\eta=0.5$	$\eta=1$	$\eta=2$	$\eta=10$
Split-plot	WP	0.043	0.057	0.054	0.052	0.058
	SP	0.053	0.062	0.069	0.075	0.088
	WP*SP	0.051	0.052	0.061	0.064	0.080
Split-plot/satterth	WP	0.038	0.056	0.054	0.052	0.058
	SP	0.052	0.062	0.069	0.075	0.088
	WP*SP	0.050	0.052	0.061	0.064	0.080
CS	WP	0.050	0.057	0.054	0.052	0.058
	SP	0.048	0.061	0.069	0.075	0.088
	WP*SP	0.049	0.052	0.061	0.064	0.080
CS/satterth	WP	0.050	0.057	0.054	0.052	0.058
	SP	0.048	0.061	0.069	0.075	0.088
	WP*SP	0.049	0.052	0.061	0.064	0.080
AR(1)	WP	0.049	0.055	0.050	0.051	0.055
	SP	0.050	0.046	0.058	0.051	0.048
	WP*SP	0.047	0.041	0.050	0.047	0.047
AR(1)/satterth	WP	0.041	0.051	0.044	0.049	0.054
	SP	0.053	0.048	0.059	0.051	0.048
	WP*SP	0.049	0.044	0.052	0.049	0.047
Percent of times AIC of AR(1)<AIC of [2]SP		44.25%	86.75%	96.40%	98.10%	99.65%

[1]a is the whole plot treatment level, b is the split-plot treatment level, r is number of replications for whole plot units, n is the sample size, and N is repetitions.

[2]SP is the conventional split-plot method. AIC of SP is exactly the same as AIC of CS model.

[3]Source: WP is whole plot main effect, SP is subplot main effect, WP*SP is interaction effect.

Table A.18 Comparison of Type I error rate and AIC value of CS and AR(1) covariance mixed model methods, and conventional split-plot method for datasets with true AR(1) covariance structure and where $a=2$, $b=9$, $r=9$, $n=162$, and $N=2000$.

Method	[3]Source	[1]$a=2$; $b=9$; $r=9$; $n=162$; $N=2000$				
		$\eta=0.1$	$\eta=0.5$	$\eta=1$	$\eta=2$	$\eta=10$
Split-plot	WP	0.047	0.050	0.049	0.046	0.043
	SP	0.056	0.061	0.084	0.087	0.123
	WP*SP	0.050	0.062	0.076	0.097	0.118
Split-plot/satterth	WP	0.042	0.048	0.049	0.046	0.043
	SP	0.056	0.061	0.084	0.087	0.123
	WP*SP	0.049	0.062	0.076	0.097	0.118
CS	WP	0.052	0.051	0.049	0.046	0.043
	SP	0.053	0.061	0.084	0.087	0.123
	WP*SP	0.047	0.062	0.076	0.097	0.118
CS/satterth	WP	0.052	0.051	0.049	0.046	0.043
	SP	0.053	0.061	0.084	0.087	0.123
	WP*SP	0.047	0.062	0.076	0.097	0.118
AR(1)	WP	0.053	0.048	0.051	0.043	0.043
	SP	0.052	0.050	0.050	0.051	0.050
	WP*SP	0.047	0.052	0.048	0.053	0.050
AR(1)/satterth	WP	0.042	0.037	0.044	0.038	0.042
	SP	0.056	0.052	0.054	0.052	0.050
	WP*SP	0.051	0.056	0.050	0.055	0.050
Percent of times AIC of AR(1)<AIC of [2]SP		49.50%	96.65%	99.30%	99.95%	100.00%

[1]a is the whole plot treatment level, b is the split-plot treatment level, r is number of replications for whole plot units, n is the sample size, and N is repetitions.

[2]SP is the conventional split-plot method. AIC of SP is exactly the same as AIC of CS model.

[3]Source: WP is whole plot main effect, SP is subplot main effect, WP*SP is interaction effect.

Appendix B - R and SAS Programs for Case 1

B.1 R Program for Case 1

####

Generate datasets with WPTrt=2; SPTrt=3; Rep=3; N=2000.

By changing the variances of u and e, we can get different variance ratio η.

##################

```
datagen<-function(x){
repli=factor(rep(1:3,each=6))
wp=factor(rep(rep(1:2,each=3),3))
sp=rep(seq(8,12,by=2),6)
Z1=model.matrix(~0+repli)
Z2=model.matrix(~0+wp:repli)
Z=cbind(Z1,Z2)
u=rnorm(9,0,1)
e=rnorm(18,0,10)
y=Z%*%u+e
y=round(y,1)
sim=data.frame(repli,wp,sp,y)}
n.2000 <- NULL
for(i in 1:2000) {
    dd <- 1:2000
    n.2000 <- rbind(n.2000, datagen(dd))
  }
sim<-factor(rep(1:2000,each=18))
data1<-cbind(n.2000,sim)

library(foreign)
write.foreign(data1, "D:/my directory/data1.txt", "D:/my   directory/data1.sas",package="SAS")
```

####

Generate datasets with WPTrt=2; SPTrt=6; Rep=3; N=2000.

By changing the variances of u and e, we can get five different variance ratio η we want.

##################

```
datagen<-function(x){
repli=factor(rep(1:3,each=12))
wp=factor(rep(rep(1:2,each=6),3))
sp=rep(seq(2,12,by=2),6)
Z1=model.matrix(~0+repli)
Z2=model.matrix(~0+wp:repli)
Z=cbind(Z1,Z2)
u=rnorm(9,0,1)
e=rnorm(36,0,10)
y=Z%*%u+e
y=round(y,1)
sim=data.frame(repli,wp,sp,y)}


n.2000 <- NULL
for(i in 1:2000) {
    dd <- 1:2000
    n.2000 <- rbind(n.2000, datagen(dd))
  }
sim<-factor(rep(1:2000,each=36))
data6<-cbind(n.2000,sim)


library(foreign)
write.foreign(data6, "E:/my directory/data6.txt", "E:/my directory/data6.sas",package="SAS")
```

```
####
Generate datasets with WPTrt=2; SPTrt=9; Rep=3; N=2000.
By changing the variances of u and e, we can get different variance ratio η.
##################
datagen<-function(x){
repli=factor(rep(1:3,each=18))
wp=factor(rep(rep(1:2,each=9),3))
sp=rep(seq(2,18,by=2),6)
Z1=model.matrix(~0+repli)
Z2=model.matrix(~0+wp:repli)
Z=cbind(Z1,Z2)
u=rnorm(9,0,1)
e=rnorm(54,0,10)
y=Z%*%u+e
y=round(y,1)
sim=data.frame(repli,wp,sp,y)}


n.2000 <- NULL
for(i in 1:2000) {
    dd <- 1:2000
    n.2000 <- rbind(n.2000, datagen(dd))
  }
sim<-factor(rep(1:2000,each=54))
data11<-cbind(n.2000,sim)


library(foreign)
write.foreign(data11, "D:/my directory/data11.txt", "D:/my
directory/data11.sas",package="SAS")
```

```
####
Generate datasets with WPTrt=2; SPTrt=3; Rep=6; N=2000.
By changing the variances of u and e, we can get different variance ratio η.
##################
datagen<-function(x){
repli=factor(rep(1:6,each=6))
wp=factor(rep(rep(1:2,each=3),6))
sp=rep(seq(8,12,by=2),12)
Z1=model.matrix(~0+repli)
Z2=model.matrix(~0+wp:repli)
Z=cbind(Z1,Z2)
u=rnorm(18,0,1)
e=rnorm(36,0,10)
y=Z%*%u+e
y=round(y,1)
sim=data.frame(repli,wp,sp,y)}

n.2000 <- NULL
for(i in 1:2000) {
    dd <- 1:2000
    n.2000 <- rbind(n.2000, datagen(dd))
  }
sim<-factor(rep(1:2000,each=36))
data16<-cbind(n.2000,sim)

library(foreign)
write.foreign(data16, "D:/my directory/data16.txt", "D:/my
directory/data16.sas",package="SAS")
```

####

Generate datasets with WPTrt=2; SPTrt=6; Rep=6; N=2000.

By changing the variances of u and e, we can get different variance ratio η.

#################

```
datagen<-function(x){
repli=factor(rep(1:6,each=12))
wp=factor(rep(rep(1:2,each=6),6))
sp=rep(seq(2,12,by=2),12)
Z1=model.matrix(~0+repli)
Z2=model.matrix(~0+wp:repli)
Z=cbind(Z1,Z2)
u=rnorm(18,0,1)
e=rnorm(72,0,10)
y=Z%*%u+e
y=round(y,1)
sim=data.frame(repli,wp,sp,y)}

n.2000 <- NULL
for(i in 1:2000) {
    dd <- 1:2000
    n.2000 <- rbind(n.2000, datagen(dd))
  }
sim<-factor(rep(1:2000,each=72))
data21<-cbind(n.2000,sim)

library(foreign)
write.foreign(data21, "D:/my directory/data21.txt", "D:/my
directory/data21.sas",package="SAS")
```

####

Generate datasets with WPTrt=2; SPTrt=9; Rep=6; N=2000.

By changing the variances of u and e, we can get different variance ratio η.

#################

```
datagen<-function(x){
repli=factor(rep(1:6,each=18))
wp=factor(rep(rep(1:2,each=9),6))
sp=rep(seq(2,18,by=2),12)
Z1=model.matrix(~0+repli)
Z2=model.matrix(~0+wp:repli)
Z=cbind(Z1,Z2)
u=rnorm(18,0,1)
e=rnorm(108,0,10)
y=Z%*%u+e
y=round(y,1)
sim=data.frame(repli,wp,sp,y)}


n.2000 <- NULL
for(i in 1:2000) {
    dd <- 1:2000
    n.2000 <- rbind(n.2000, datagen(dd))
  }
sim<-factor(rep(1:2000,each=108))
data26<-cbind(n.2000,sim)


library(foreign)
write.foreign(data26, "D:/my directory/data26.txt", "D:/my
directory/data26.sas",package="SAS")
```

```
####
```

Generate datasets with WPTrt=2; SPTrt=3; Rep=9; N=2000.

By changing the variances of u and e, we can get different variance ratio η.

```
##################
datagen<-function(x){
repli=factor(rep(1:9,each=6))
wp=factor(rep(rep(1:2,each=3),9))
sp=rep(seq(8,12,by=2),18)
Z1=model.matrix(~0+repli)
Z2=model.matrix(~0+wp:repli)
Z=cbind(Z1,Z2)
u=rnorm(27,0,1)
e=rnorm(54,0,10)
y=Z%*%u+e
y=round(y,1)
sim=data.frame(repli,wp,sp,y)}


n.2000 <- NULL
for(i in 1:2000) {
    dd <- 1:2000
    n.2000 <- rbind(n.2000, datagen(dd))
  }
sim<-factor(rep(1:2000,each=54))
data31<-cbind(n.2000,sim)


library(foreign)
write.foreign(data31, "D:/my directory/data31.txt", "D:/my
directory/data31.sas",package="SAS")
```

```
####
Generate datasets with WPTrt=2; SPTrt=6; Rep=9; N=2000.
By changing the variances of u and e, we can get different variance ratio η.
##################
datagen<-function(x){
repli=factor(rep(1:9,each=12))
wp=factor(rep(rep(1:2,each=6),9))
sp=rep(seq(2,12,by=2),18)
Z1=model.matrix(~0+repli)
Z2=model.matrix(~0+wp:repli)
Z=cbind(Z1,Z2)
u=rnorm(27,0,1)
e=rnorm(108,0,10)
y=Z%*%u+e
y=round(y,1)
sim=data.frame(repli,wp,sp,y)}

n.2000 <- NULL
for(i in 1:2000) {
    dd <- 1:2000
    n.2000 <- rbind(n.2000, datagen(dd))
  }
sim<-factor(rep(1:2000,each=108))
data36<-cbind(n.2000,sim)

library(foreign)
write.foreign(data36, "D:/my directory/data36.txt", "D:/my
directory/data36.sas",package="SAS")
```

```
####
Generate datasets with WPTrt=2; SPTrt=9; Rep=9; N=2000.
By changing the variances of u and e, we can get different variance ratio η.
##################
datagen<-function(x){
repli=factor(rep(1:9,each=18))
wp=factor(rep(rep(1:2,each=9),9))
sp=rep(seq(2,18,by=2),18)
Z1=model.matrix(~0+repli)
Z2=model.matrix(~0+wp:repli)
Z=cbind(Z1,Z2)
u=rnorm(27,0,1)
e=rnorm(162,0,10)
y=Z%*%u+e
y=round(y,1)
sim=data.frame(repli,wp,sp,y)}

n.2000 <- NULL
for(i in 1:2000) {
    dd <- 1:2000
    n.2000 <- rbind(n.2000, datagen(dd))
  }
sim<-factor(rep(1:2000,each=162))
data41<-cbind(n.2000,sim)

library(foreign)
write.foreign(data41, "D:/my directory/data41.txt", "D:/my
directory/data41.sas",package="SAS")
```

B.2 SAS Program for Case 1

```
/*Input the dataset been generated through R program*/
DATA  rdata ;
INFILE  "D:/my directory/data1.txt"
    DSD
    LRECL= 21 ;
INPUT
 repli
 wp
 sp
 y
 sim
;
FORMAT repli repli. ;
FORMAT wp wp. ;
FORMAT sim sim. ;
RUN;

/*conventional split plot model*/
/*By deleting the "/ddfm=satterth", we get the default way */
proc mixed data=rdata ic;
  by sim;
  class repli wp sp;
  model y=wp sp wp*sp/ddfm=satterth;
  random repli(wp);
  lsmeans wp sp wp*sp;
  ods output Tests3=test0 infocrit=critical;
run;

/*Check type 1 error rate for geno, fert, and geno*fert*/
data type1wp0;set test0;
  by sim;
  where ProbF<0.05 and Effect="wp";
run;
data type1sp0;set test0;
  by sim;
  where ProbF<0.05 and Effect="sp";
run;
data type1wp_sp0;set test0;
  by sim;
  where ProbF<0.05 and Effect="wp*sp";
run;
```

```
/*CS covariance structure model*/
/*By deleting the "/ddfm=satterth", we get the default way */
proc mixed data=rdata ic;
  by sim;
  class repli wp sp;
  model y=wp sp wp*sp/ddfm=satterth;
  repeated sp/subject=repli(wp) type=cs r rcorr;
  lsmeans wp sp wp*sp;
  ods output Tests3=test1 infocrit=critical1;
run;
data type1wp1;set test1;
  by sim;
  where ProbF<0.05 and Effect="wp";
run;
data type1sp1;set test1;
  by sim;
  where ProbF<0.05 and Effect="sp";
run;
data type1wp_sp1;set test1;
  by sim;
  where ProbF<0.05 and Effect="wp*sp";
run;

/*AR(1) covariance structure model*/
/*By deleting the "/ddfm=satterth", we get the default way */
proc mixed data=rdata ic;
  by sim;
  class repli wp sp;
  model y=wp sp wp*sp/ddfm=satterth;
  repeated sp/subject=repli(wp) type=ar(1) r rcorr;
  lsmeans wp sp wp*sp;
  ods output Tests3=test2 infocrit=critical2;
run;
data type1wp2;set test2;
  by sim;
  where ProbF<0.05 and Effect="wp";
run;
data type1sp2;set test2;
  by sim;
  where ProbF<0.05 and Effect="sp";
run;
data type1wp_sp2;set test2;
  by sim;
  where ProbF<0.05 and Effect="wp*sp";
run;
```

```
/*Sort data to count the times that the AIC values of conventional split-plot model is smaller than
that of the AR(1) covariance model*/
proc sort data=Critical(keep=AIC sim);
  by sim;
run;
proc sort data=Critical2(keep=AIC sim);
  by sim;
run;
data null;
  retain count 0;
  merge Critical
     Critical2(rename=(AIC=AICar1));
  by sim;
  if AIC<AICar1
    then do;
    count +1;
  end;
run;
```

Appendix C - SAS Programs for Case 2

```
/* Generate datasets with WPTrt=2; SPTrt=3; Rep=3; N=2000; rho=1/11.*/
/* By changing values of i and k, we can change the values of SPtrt and Rep */
/* In this program, the value of rho can be changed corresponding to the value of η in Case 1*/
/* In this program, the value of sigma is chosen corresponding to variances of u and e in Case 1*/

%macro Simulate(Rounds);
%do m = 1 %to &Rounds;
data temp&m;
  do i = 1 to 3;
    do j=1 to 2;
        rho=1/11;
        sigma = rannor(0)*sqrt(11);
          do k = 1 to 3;
              day = k;
              if k = 1 then
                  do; sigma = rannor(0)*sqrt(11); s = 0; output; end;
               else do;
                  s = rannor(0)*sqrt((1-rho*rho)*11);
                  sigma = (rho)*sigma+s;
                  output;
              end;
        end;
    end;
  end;
run;
data temp&m(keep=i j day round y);
  set temp&m;
  round=&m;
  y = sigma;
run;
%end;
%mend;
%Simulate(2000);run;
data ardata;
 set Temp1-Temp2000;
run;
data ardata;
 rename i=repli
     j=wp
     day=sp
     round=sim;
 set ardata;
run;
```

```
/*conventional split plot model*/
/*By deleting the "/ddfm=satterth", we get the default way */
proc mixed data=rdata ic;
  by sim;
  class repli wp sp;
  model y=wp sp wp*sp/ddfm=satterth;
  random repli(wp);
  lsmeans wp sp wp*sp;
  ods output Tests3=test0 infocrit=critical;
run;
/*Check type 1 error rate for geno, fert, and geno*fert*/
data type1wp0;set test0;
  by sim;
  where ProbF<0.05 and Effect="wp";
run;
data type1sp0;set test0;
  by sim;
  where ProbF<0.05 and Effect="sp";
run;
data type1wp_sp0;set test0;
  by sim;
  where ProbF<0.05 and Effect="wp*sp";
run;

/*CS covariance structure model*/
/*By deleting the "/ddfm=satterth", we get the default way */
proc mixed data=rdata ic;
  by sim;
  class repli wp sp;
  model y=wp sp wp*sp/ddfm=satterth;
  repeated sp/subject=repli(wp) type=cs r rcorr;
  lsmeans wp sp wp*sp;
  ods output Tests3=test1 infocrit=critical1;
run;
data type1wp1;set test1;
  by sim;
  where ProbF<0.05 and Effect="wp";
run;
data type1sp1;set test1;
  by sim;
  where ProbF<0.05 and Effect="sp";
run;
data type1wp_sp1;set test1;
  by sim;
  where ProbF<0.05 and Effect="wp*sp";
run;
```

```
/*AR(1) covariance structure model*/
/*By deleting the "/ddfm=satterth", we get the default way */
proc mixed data=rdata ic;
  by sim;
  class repli wp sp;
  model y=wp sp wp*sp/ddfm=satterth;
  repeated sp/subject=repli(wp) type=ar(1) r rcorr;
  lsmeans wp sp wp*sp;
  ods output Tests3=test2 infocrit=critical2;
run;
data type1wp2;set test2;
  by sim;
  where ProbF<0.05 and Effect="wp";
run;
data type1sp2;set test2;
  by sim;
  where ProbF<0.05 and Effect="sp";
run;
data type1wp_sp2;set test2;
  by sim;
  where ProbF<0.05 and Effect="wp*sp";
run;

/*Sort data to count the times that the AIC values of conventional split-plot model is larger than
that of the AR(1) covariance model*/
proc sort data=Critical(keep=AIC sim);
  by sim;
run;
proc sort data=Critical2(keep=AIC sim);
  by sim;
run;
data null;
  retain count 0;
  merge Critical
      Critical2(rename=(AIC=AICar1));
  by sim;
  if AIC>AICar1
  then do;
  count +1;
 end;
run;
```