MODEL SELECTION TECHNIQUES FOR REPEATED MEASURES COVARIANCE STRUCTURES

E. Barry Moser

Raul E. Macchiavelli

Recommended Citation

Moser, E. Barry and Macchiavelli, Raul E. (2002). "MODEL SELECTION TECHNIQUES FOR REPEATED MEASURES COVARIANCE STRUCTURES," *Conference on Applied Statistics in Agriculture*. https://doi.org/10.4148/2475-7772.1197

MODEL SELECTION TECHNIQUES FOR REPEATED MEASURES COVARIANCE STRUCTURES

E. Barry Moser, Dept. of Experimental Statistics, Louisiana State Univ. AgCenter, Baton Rouge, LA 70803-5606
and
Raúl E. Macchiavelli, Dept. of Agronomy and Soils, Univ. of Puerto Rico, P.O. Box 9030, Mayagüez, PR 00681-9030

Abstract

A parsimonious covariance structure of repeated measures is often sought for purposes of increased power for testing hypotheses about the means, and for insight into the stochastic processes governing the repeated measures. For normal data, model selection is often based upon likelihood ratio tests or information criteria derived from the likelihood, sometimes supplemented with graphical plots of correlations and partial correlations. We exploit the ordered nature of repeated measures to decompose the likelihood ratio goodness-of-fit test statistic, and display graphical fingerprints associated with the covariance structures to help detect covariance structure misspecification, in order to provide guidance in choosing an appropriate structure for the data. The proposed methodology is illustrated with simulated repeated measures data and then applied to an experiment to compare tillage methods of pasture establishment.

1. Introduction

Repeated measurements taken on the same unit have dependencies that must be properly accounted for in their analysis. In some cases, an unstructured multivariate approach may be used, though it is often desirable to use a parsimonious covariance structure to describe the associations among the repeated measures for purposes of increased power for testing hypotheses about the means, and for insight into the stochastic processes that govern the repeated measures. In normal linear models, this is attained by assuming a particular structure for the covariance matrix (Dempster 1972, Lindsey 1993, Jennrich and Schluchter 1986, Wolfinger 1993, Macchiavelli and Moser 1996, Wolfinger 1996). The covariance structure is often regarded as a nuisance parameter, so that the ideal solution is to use a parsimonious covariance model that describes well the observed data. If an unrealistically simple covariance structure is selected, the tests will be invalid. On the other hand, an unnecessarily complex structure (or no structure at all) decreases the power (Macchiavelli and Moser 1997).

Wolfinger (1993) mentions several steps involved in making inferences about the mean parameters in these models. The first step is to fit a mean model (overfitting may be necessary to avoid creating spurious dependencies), and then initial covariance structures can be studied using relevant theory and graphics. Finally formal statistical techniques can be applied to select the final covariance structure and make inferences for the mean parameters. From a practical perspective, this strategy has two aspects that are critical to ensure a good covariance structure is chosen and fit: the selection of the structure and the likelihood optimization problem for fitting the selected structure. Methods for choosing appropriate structures are mostly based on model

selection techniques such as a sequence of likelihood ratio tests (Lindsey 1993, Diggle et al. 1994) or penalized likelihood criteria (Macchiavelli and Arnold 1995). In the related area of linear structural equations (Jöreskog and Sörbom 1988), indices of goodness of fit and residuals have been developed (Browne 1982, SAS 1996).

We exploit the ordered nature of the repeated measures along with decompositions of the covariance matrix to generate partitions of the likelihood ratio test statistic and of graphics to produce diagnostics that help isolate and highlight specific problems in the lack of fit of a covariance structure. Further, these methods can give guidance in proposing appropriate structures to consider. First, we will show how the likelihood ratio test statistic can be partitioned, and then we'll show how specific covariance structures can be compared graphically. Finally, we'll illustrate these methods with both simulated and real examples. The methods that we are proposing are not replacements for existing methodologies, such as the PRISM plots of Zimmerman (2000), but are to be used along with these methods.

2. Partition of the Likelihood Ratio Test Statistic

Let

$$\mathbf{Y} = \left(Y_1, ..., Y_T \right) \sim N_T \left(\boldsymbol{\mu}, \boldsymbol{\Sigma} \right) \tag{2.1}$$

be a T-dimensional vector representing repeated observations on a single experimental unit. Consider n independent random vectors $\left(\mathbf{Y}_1, ..., \mathbf{Y}_n \right)$, each with the same covariance matrix. Define

$$\mathbf{Y}_{n \times T} = \begin{pmatrix} \mathbf{Y}_1' \\ ... \\ \mathbf{Y}_n' \end{pmatrix} \tag{2.2}$$

and assume a linear model $E(\mathbf{Y}_{n \times T}) = \mathbf{X}_{n \times a} \boldsymbol{\Theta}_{a \times T}$, for some $n \times a$ matrix $\mathbf{X}$ of full column rank. The elements of $\boldsymbol{\Theta}$ are functionally independent.

We are usually interested in making inferences about $\boldsymbol{\Theta}$, the mean parameters, while the covariance matrix, $\boldsymbol{\Sigma}$, contains the nuisance parameters. Since the multivariate normal distribution is an exponential family, the canonical parameterization for the nuisance parameter is not in terms of $\boldsymbol{\Sigma}$ but in terms of the concentration matrix, $\boldsymbol{\Omega} = \boldsymbol{\Sigma}^{-1}$. The likelihood for n independent observations using this parameterization is:

$$\mathsf{L}(\mathbf{Y}; \boldsymbol{\Theta}, \boldsymbol{\Omega}) = (2\pi)^{-nT/2} |\boldsymbol{\Omega}|^{n/2} \exp \left\{ -\tfrac{1}{2} \mathrm{tr} \left[\boldsymbol{\Omega} \sum_{i=1}^{n} \left(\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\Theta} \right) \left(\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\Theta} \right)' \right] \right\}, \tag{2.3}$$

where $\mathbf{X}_i$ is the ith row of $\mathbf{X}$.

As mentioned before, several structures are commonly used for $\boldsymbol{\Omega} = (\omega_{ij})$ or $\boldsymbol{\Sigma} = (\sigma_{ij})$. The compound symmetry structure assumes $\sigma_{ii} = a + b$ and $\sigma_{ij} = b$ if $i \neq j$. The first order autoregressive model assumes the $\sigma_{ij} = \sigma \rho^{|i-j|}$. The first order antedependence structure assumes that $\omega_{ij} = 0$ if $|i - j| > 1$. In the particular case of structures defined only by conditional

Applied Statistics in Agriculture

independence (zeroes in Ω), we have a linear exponential family (Dempster 1972, Cox and Wermuth 1990). From the likelihood function, it is clear that the MLE of Σ and Ω for complete data if no structure is assumed are, respectively, $\hat{\Sigma}_U = S$ and $\hat{\Omega}_U = S^{-1}$, where

$$S = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - X_i\hat{\Theta}\right)\left(Y_i - X_i\hat{\Theta}\right)'. \tag{2.4}$$

Under a particular structure S, the MLEs of Σ and Ω will be denoted as $\hat{\Sigma}_S$ and $\hat{\Omega}_S$.

In order to test the hypothesis that the structure proposed is valid against a general unstructured alternative, the likelihood ratio chi-square test statistic can be written as

$$\chi^2 = -2\log\Lambda = n\log\left|\hat{\Sigma}_S\hat{\Sigma}_U^{-1}\right| + \sum_{i=1}^{n}\underline{e}_{iS}'\hat{\Omega}_S\underline{e}_{iS} - \sum_{i=1}^{n}\underline{e}_{iU}'\hat{\Omega}_U\underline{e}_{iU}, \tag{2.5}$$

where $\underline{e}_{iS} = \left(Y_i - X_i\hat{\Theta}_S\right)$ and $\underline{e}_{iU} = \left(Y_i - X_i\hat{\Theta}_U\right)$ are the residuals under the proposed structure and the unstructured matrices respectively.

Using spectral decompositions of the covariance matrix we can write

$$\hat{\Sigma}_U^{-1} = \hat{D}_U^{-1/2}\left[\sum_{j=1}^{T}\frac{1}{\lambda_{jU}}P_{jU}P_{jU}'\right]\hat{D}_U^{-1/2} \tag{2.6}$$

and

$$\underline{e}_{iU}'\hat{\Omega}_U\underline{e}_{iU} = \sum_{j=1}^{T}\frac{\left(\underline{e}_{iU}'\hat{D}_U^{-1/2}P_{jU}\right)^2}{\lambda_{jU}}. \tag{2.7}$$

Here $\hat{D}_U$ is the matrix containing the diagonal elements of $\hat{\Sigma}_U$, and λ_{jU} and P_{jU} are the jth eigenvalue and eigenvector respectively of $\hat{R}_U = \hat{D}_U^{-1/2}\hat{\Sigma}_U\hat{D}_U^{-1/2}$, the correlation matrix.

A similar decomposition can be obtained for $\hat{\Sigma}_S$. Using these decompositions, the likelihood ratio chi-square test statistic can be written as

$$\chi^2 = \sum_{i=1}^{n}\sum_{j=1}^{T}d_{ij}, \tag{2.8}$$

where

$$d_{ij} = \log\lambda_{jS} + \tfrac{1}{T}\log\left|\hat{D}_S\right| + \frac{\left(\underline{e}_{iS}'\hat{D}_S^{-1/2}P_{jS}\right)^2}{\lambda_{jS}}$$

$$-\left(\log\lambda_{jU} + \tfrac{1}{T}\log\left|\hat{D}_U\right| + \frac{\left(\underline{e}_{iU}'\hat{D}_U^{-1/2}P_{jU}\right)^2}{\lambda_{jU}}\right).$$

In order to have good estimates of the unstructured matrices, the residuals $\underline{e}_i$ need to be unbiased. Hence it is better to have a mean "saturated" model (Wolfinger 1993).

Since (2.8) has an asymptotic expected value of df (equal to the difference in the number of parameters in the proposed structured and in the unstructured covariances), we can divide each component d_{ij} by df to spot individual contributions that are too large. Other ways of looking at the contributions of these components is by adding them up by subjects or by dimensions. Thus, $\sum_{j=1}^{T} d_{ij}$ or $\sum_{j=1}^{T} d_{ij}\Big/ df$ will indicate if a particular subject is contributing too much to the lack of fit of the proposed structure. Similarly, $\sum_{i=1}^{n} d_{ij}$ or $\sum_{i=1}^{n} d_{ij}\Big/ df$ will indicate if certain dimensions are causing the lack of fit.

From (2.5) another decomposition of the test statistic can reflect the aspect of the model in which there is lack of fit. By substituting $\hat{\Sigma}_S = \hat{D}_S^{1/2} \hat{R}_S \hat{D}_S^{1/2}$ and $\hat{\Sigma}_U^{-1} = \hat{D}_U^{-1/2} \hat{R}_U^{-1} \hat{D}_U^{-1/2}$ in (2.5), we can write it as

$$\chi^2 = n \log\left|\hat{D}_S \hat{D}_U^{-1}\right| + n \log\left|\hat{R}_S \hat{R}_U^{-1}\right| + \sum_{i=1}^{n} \left(\underline{e}'_{iS} \hat{\Omega}_S \underline{e}_{iS} - \underline{e}'_{iU} \hat{\Omega}_U \underline{e}_{iU} \right). \tag{2.9}$$

We can see that (2.9) can be divided into three terms: one associated with the variance structure, one associated with the correlation structure and one associated with the difference in the residuals. As before, the contribution of each of these terms to the lack of fit can be studied dividing each term by the degrees of freedom: a component much larger than 1 would indicate that that particular aspect of the structure is contributing the most to the lack of fit. The terms associated with the variance and correlation parts can further be investigated by computing ratios and log-ratios of corresponding variances in $\hat{D}_S$ and $\hat{D}_U$, and by computing ratios and log-ratios of corresponding eigenvalues of $\hat{R}_S$ and $\hat{R}_U$. Note that the component contributions to the likelihood ratio statistic are simply n times the component log-ratios. Thus (2.9) can be re-written as

$$\chi^2 = \sum_{i=1}^{T} n \log(\hat{\sigma}_{iS}^2 / \hat{\sigma}_{iU}^2) + \sum_{i=1}^{T} n \log(\hat{\lambda}_{iS} / \hat{\lambda}_{iU}) + \sum_{i=1}^{n} \left(\underline{e}'_{iS} \hat{\Omega}_S \underline{e}_{iS} - \underline{e}'_{iU} \hat{\Omega}_U \underline{e}_{iU} \right) \tag{2.10}$$

where $\hat{\sigma}_{iS}^2$ and $\hat{\sigma}_{iU}^2$ are the variances from $\hat{\Sigma}_S$ and $\hat{\Sigma}_U$, respectively, and $\hat{\lambda}_{iS}$ and $\hat{\lambda}_{iU}$ are the eigenvalues of $\hat{R}_S$ and $\hat{R}_U$, respectively.

Applied Statistics in Agriculture

3. Graphical Fingerprints

The eigenstructures of several often-used models for repeated measures are well known, and when the time variables are displayed graphically in the space of the eigenvalues, produce characteristic plots. For the compound symmetry structure, here given for 4 times,

$$\Sigma = \begin{bmatrix} \alpha + \beta & \beta & \beta & \beta \\ \beta & \alpha + \beta & \beta & \beta \\ \beta & \beta & \alpha + \beta & \beta \\ \beta & \beta & \beta & \alpha + \beta \end{bmatrix} \tag{3.1}$$

there are two, usually unique, eigenvalues, λ_1 and $\lambda_{i \neq 1}$, and the set of eigenvectors, **E**, associated with the smaller of the two describe a sphere where

$$\lambda = \begin{bmatrix} \alpha + 4\beta \\ \alpha \\ \alpha \\ \alpha \end{bmatrix} \text{ and } \mathbf{E'} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} \tag{3.2}$$

and where the eigenvectors are the rows of **E′**. When the time variables are plotted in this space, the time vectors are orthogonal and describe a sphere that indicates equal correlations among any pair of time vectors.

Consider a first-order autoregressive structure for 4 times with covariance matrix

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}. \tag{3.3}$$

The eigenvalues are

$$\lambda = \begin{bmatrix} \dfrac{1}{2}\rho^2 - \dfrac{1}{2}\rho + 1 + \dfrac{1}{2}\sigma^2(\rho+1)\sqrt{\rho^4 - 2\rho^3 + 5\rho^2} \\ \dfrac{1}{2}\rho^2 - \dfrac{1}{2}\rho + 1 - \dfrac{1}{2}\sigma^2(\rho+1)\sqrt{\rho^4 - 2\rho^3 + 5\rho^2} \\ -\dfrac{1}{2}\rho^2 - \dfrac{1}{2}\rho - 1 + \dfrac{1}{2}\sigma^2(\rho-1)\sqrt{\rho^4 + 2\rho^3 + 5\rho^2} \\ -\dfrac{1}{2}\rho^2 - \dfrac{1}{2}\rho - 1 - \dfrac{1}{2}\sigma^2(\rho-1)\sqrt{\rho^4 + 2\rho^3 + 5\rho^2} \end{bmatrix} \tag{3.4}$$

and the associated eigenvectors (in row order) are

$$\mathbf{E}' = \begin{bmatrix} -\dfrac{\sigma^2(\rho+1)-\lambda_1}{\sigma^2\rho(\rho+1)} & 1 & 1 & -\dfrac{\sigma^2(\rho+1)-\lambda_1}{\sigma^2\rho(\rho+1)} \\[2mm] -\dfrac{\sigma^2(\rho+1)-\lambda_2}{\sigma^2\rho(\rho+1)} & 1 & 1 & -\dfrac{\sigma^2(\rho+1)-\lambda_2}{\sigma^2\rho(\rho+1)} \\[2mm] -1 & \dfrac{\sigma^2(\rho^3-1)+\lambda_3}{\sigma^2\rho(\rho-1)} & -\dfrac{\sigma^2(\rho^3-1)+\lambda_3}{\sigma^2\rho(\rho-1)} & 1 \\[2mm] -1 & \dfrac{\sigma^2(\rho^3-1)+\lambda_4}{\sigma^2\rho(\rho-1)} & -\dfrac{\sigma^2(\rho^3-1)+\lambda_4}{\sigma^2\rho(\rho-1)} & 1 \end{bmatrix} \quad (3.5)$$

where the λ_i are the elements of $\boldsymbol{\lambda}$. For a larger number of time periods, such symmetry persists among the eigenvalues and eigenvectors. For example, consider a first-order autoregressive process with 6 times. When the time vectors are plotted in the first two dimensions corresponding with the 2 largest eigenvalues, the structure gives a very symmetric pattern that sequentially orders the times (Figure 1). When the time vectors are plotted in the space of the last two eigenvectors, a star pattern results (Figure 2).

Other common covariance structures also have specific patterns, though the actual pattern can depend considerably upon the values of the parameters for the structure. We propose that the time vectors from the unstructured covariance model and from the proposed (reduced) covariance model be plotted in the space of the eigenvectors of the proposed model. In this way, one can visually examine the goodness-of-fit of the proposed model to the unstructured model. In addition, substantial deviations of fit can often suggest reasons for the lack-of-fit. Since the variances can have substantial effects on these results, we propose that the fingerprint graphics be performed on the correlation matrices, with separate diagnostics used for modeling the variances.

A scaling of the eigenvectors using their corresponding eigenvalues is performed on each structure, then a Procrustes rotation of the unstructured covariance matrix time vectors is used to most closely match the proposed structure before the graphics are produced. Let $\mathbf{E}_U$ and $\mathbf{E}_S$ be the scaled eigenvectors for the unstructured and structured models, respectively, then

$$\mathbf{E}_U'\mathbf{E}_S = \mathbf{UDV}' \qquad (3.6)$$

using the singular value decomposition, so that the Procrustes rotation of $\mathbf{E}_U$ to $\hat{\mathbf{E}}_U$ to most closely match $\mathbf{E}_S$ is

$$\hat{\mathbf{E}}_U = \mathbf{E}_S\mathbf{UV}'. \qquad (3.7)$$

In addition, the Euclidean distance from one structure to the other is computed following the rotation.

Applied Statistics in Agriculture

4. Examples

To illustrate the methodologies, a simulated data set having a heterogeneous first-order autoregressive structure was generated and analyzed. The analysis considered a variety of covariance structures available through PROC MIXED of SAS (SAS Institute, Inc., Cary, NC). The simulation illustration is followed by the analysis of a pasture forage experiment involving repeated measures.

4.1 Simulated examples

The simulated data set of 6 time equally-spaced time periods was generated under a heterogeneous first-order autoregressive process (ARH(1)), and then a variety of covariance structure models were fit to these data. The ARH(1) model fits adequately ($X^2 = 19.8$, df=14, P=0.14), but more interesting are the diagnostics associated with structures such as CS, CSH, and AR(1). The CS structure should not be a good fit to either the variances or covariances of this data set. This is verified both by the likelihood ratio diagnostics (Table 1), where the contributions to the likelihood ratio test are very large for both the variance and covariance components, and the fingerprint graphics (Figure 3), where the structure of the correlations is not well modeled by the CS structure. For the CSH model, the variances of the data are well modeled, the likelihood ratio decomposition (Table 2) has a small contribution to the variances and still a large one for the covariances, and so the correlations are again poorly described. The AR(1) model describes well the correlations for these data, but does not handle the heterogeneous variances well with a large contribution from the variances to the likelihood ratio test (Table 3). Finally, the ARH(1) model is fit to these data and both the correlations (Figure 4) and likelihood ratio test components are all small (Table 4).

4.2 Pasture forages

In order to compare minimum and non-tillage to conventional methods of pasture establishment, 4 different treatments were tested in the central humid region of Puerto Rico. The species used was *Brachiaria decumbens* cv. Basilik (Signal Grass). The treatments were:

1. minimum tillage (one superficial harrow disk pass) followed with seeds broadcast on the plot;
2. minimum tillage, one broadcast application of herbicide, and 45 days later, minimum tillage and seeds broadcast on the tilled plot;
3. minimum tillage, one broadcast application of herbicide, and 45 days later, 2 disc passes every 2m (90° grid pattern) followed by seeds broadcast on the tilled plot;
4. non-tillage, one broadcast application of herbicide, and 45 days later, seeds broadcast on the plot;
5. conventional planting: 2 disc plow passes followed by 2 harrow disc passes then seeds broadcast on the plot.

The experiment was arranged in a randomized complete block design with 3 replicates. There were measurements of coverage (average estimated percentages of cover from 10 frames thrown at random in each plot), and plant density (average number of plants in the 10 frames).

There are 5 repeated measurements, taken at 1-month intervals from August to December 2001. In this example we analyze the measures of cover.

Models addressing heterogeneity among the variances of the times are important for these data. Consider first the heterogeneous compound symmetry (CSH) model. The likelihood ratio test clearly rejects this structure as appropriate for these data (X^2=76.6, df=9, P<0.0001) and its decomposition clearly shows that the major lack-of-fit is with the correlations (Table 5). The fingerprints for this model indicate that some ordering of the times is needed, and also shows that this model does not adequately fit (Figure 5). The heterogeneous first-order autoregressive process does not fit these data much better (X^2=70.8, df=9, P<0.0001) and also has problems fitting the correlation structure (Table 6, Figures 6-7). Notice, however, that the fingerprints in the last two principal axes suggest that the AR(1) process has captured some of the time-ordering of the time vectors (Figure 7). Although the overall contribution to the likelihood ratio test by the variances is not large (X^2=6.77), there is considerable variation among the individual comparisons suggesting some lack-of-fit (Table 7). Contributions to the lack-of-fit, however, are very great in dimension 5 comparing correlation structures (Table 8). This would suggest problems in modelling well the partial correlations. A two-factor factor analytic model (FA(2)), on the other hand does appear to model the data reasonably well (X^2=2.2, df=3, P=0.53). There are no large contributions to the likelihood ratio statistic and the observed time vectors are very near the fitted vectors in the fingerprint plot (Table 9, Figures 8-9). The decompositions of both the variances (Table 10) and correlations (Table 11) show very good fits of this model. The unstructured with two minor bands model (UN(3)) also fits these data reasonably well (X^2=0.69, df=3, P=0.88) and has diagnostics similar to the FA(2) model.

5. Summary

Both numerical and graphical decompositions of covariance structures are shown to be useful diagnostics for model selection for repeated measures covariance structures. The likelihood ratio test statistic for comparing structures with an unstructured alternative is written in terms of comparisons of variances, correlations, and residuals, and so lack-of-fit can be attributed to specific components which may suggest alternative structures to consider. In addition, the time vectors of the repeated measures are plotted in the space of the principal components of the fitted covariance structure, and the unstructured fit is projected into this same space providing a graphical assessment of fit. These tools should be complementary to existing diagnostics for repeated measures covariance structure model selection.

References

Browne, M.W. (1982) Covariance Structures. In *Topics in Multivariate Analysis*, D.M. Hawkins, ed. New York: Cambridge University Press.

Cox, D.R. and Wermuth, N. (1990) An approximation to maximum likelihood estimates in reduced models. *Biometrika* 77: 747-761.

Dempster, A.P. (1972) Covariance Selection. *Biometrics* 28: 157-175.

Diggle, P., Liang, K and Zeger, S. (1994) *Analysis of Longitudinal Data*. London: Oxford Univ. Press.

Jennrich, R. and Schluchter,M. (1986) Unbalanced repeated measures models with structured covariance matrices. *Biometrics* 42: 805-820.

Jöreskog, K. and Sörbom, D. (1988) *LISREL 7: A guide to the program and applications*. Chicago: SPSS, Inc.

Lindsey, J. (1993) *Models for repeated measurements*. London: Oxford Univ. Press.

Macchiavelli, R. and S. Arnold (1995) Difficulties with the use of penalized likelihood criteria in antedependence and polynomial models. *Communications in Statistics, Theory and Methods*, 24 (2): 501-522.

Macchiavelli, R. and E.B. Moser (1996) Tests of hypotheses in repeated measures models under general ante-dependence covariance structures in SAS software. *Proceedings of the 21st. SAS Users Group International Conference*, 1265-68.

Macchiavelli, R. and E.B. Moser (1997) Analysis of repeated measures with ante-dependence models. *Biometrical Journal* 39 (3) :339-350.

SAS Institute, Inc. (1996) *SAS/STAT Software: Changes and Enhancements through Release 6.12*. Cary, NC: SAS Institute Inc.

Wolfinger, R. (1993) Covariance structure selection in general mixed models. *Communications in Statistics, Theory and Methods*, 22(4): 1079-1106.

Wolfinger, R. (1996) Heterogeneous variance-covariance structures for repeated measures. *Journal of Agricultural, Biological, and Environmental Statistics,* 1:205-230.

Zimmerman, D.L. (2000) Viewing the correlation structure of longitudinal data through a PRISM. *The American Statistician* 54(4): 310-318.

Table 1. Likelihood ratio test (df=19) decomposition for the ARH(1) data fitted with a homogeneous compound symmetry (CS) model.

Component	Contribution to Chi-square	Contribution/df
Variances	111.8	5.9
Correlations	124.8	6.6
Residuals	0.0	0.00
Total	236.6	12.5

Table 2. Likelihood ratio test (df=14) decomposition for the ARH(1) data fitted with a heterogeneous compound symmetry (CSH) model.

Component	Contribution to Chi-square	Contribution/df
Variances	0.6	0.05
Correlations	80.5	5.75
Residuals	0.0	0.00
Total	81.1	5.80

Table 3. Likelihood ratio test (df=19) decomposition for the ARH(1) data fitted with a homogeneous first-order autoregressive (AR(1)) model.

Component	Contribution to Chi-square	Contribution/df
Variances	118.4	6.2
Correlations	44.2	2.3
Residuals	0.0	0.00
Total	162.6	8.5

Table 4. Likelihood ratio test (df=14) decomposition for the ARH(1) data fitted with a heterogeneous first-order autoregressive (ARH(1)) model.

Component	Contribution to Chi-square	Contribution/df
Variances	-2.4	-0.17
Correlations	22.2	1.59
Residuals	0.0	0.00
Total	19.8	1.42

Table 5. Likelihood ratio test (df=9) decomposition for the pasture forage data fitted with a heterogeneous compound symmetry (CS) model.

Component	Contribution to Chi-square	Contribution/df
Variances	11.5	1.3
Correlations	65.1	7.2
Residuals	0.0	0.00
Total	76.6	8.5

Applied Statistics in Agriculture

Table 6. Likelihood ratio test (df=9) decomposition for the pasture forage data fitted with a heterogeneous first-order autoregressive (ARH(1)) model.

Component	Contribution to Chi-square	Contribution/df
Variances	6.8	0.75
Correlations	64.0	7.12
Residuals	0.0	0.00
Total	70.8	7.87

Table 7. Likelihood ratio test (df=9) decomposition of variances for the pasture forage data fitted with a heterogeneous first-order autoregressive (ARH(1)) model.

Time	V(Fitted)	V(Full)	VR	Log(VR)	LR
T1	39.94384	27.938965	1.430	0.3575	5.36
T2	75.514134	40.247919	1.876	0.6293	9.44
T3	164.86238	116.72184	1.412	0.3453	5.18
T4	120.34904	172.24677	0.699	-0.3585	-5.38
T5	92.973475	156.70796	0.593	-0.5221	-7.83
Sum	.	.	.	.	6.77

Table 8. Likelihood ratio test (df=9) decomposition of eigenvalues of correlation matrix for the pasture forage data fitted with a heterogeneous first-order autoregressive (ARH(1)) model.

Dimension	E(Fitted)	E(Full)	ER	Log(ER)	LR
D1	2.668577	2.4523415	1.088	0.0845	1.27
D2	1.1511716	1.5197386	0.757	-0.2778	-4.17
D3	0.5622654	0.8462221	0.664	-0.4088	-6.13
D4	0.3496942	0.1776492	1.968	0.6772	10.16
D5	0.2682918	0.0040487	66.267	4.1937	62.91
Sum	.	.	.	.	64.03

Table 9. Likelihood ratio test (df=3) decomposition for the pasture forage data fitted with a two-factor factor analytic (FA(2)) model.

Component	Contribution to Chi-square	Contribution/df
Variances	1.1	0.37
Correlations	1.1	0.37
Residuals	0.0	0.00
Total	2.2	0.74

Table 10. Likelihood ratio test (df=3) decomposition of variances for the pasture forage data fitted with a two-factor factor analytic (FA(2)) model.

Time	V(Fitted)	V(Full)	VR	Log(VR)	LR
T1	27.355754	27.938965	0.979	-0.0211	-0.32
T2	36.894511	40.247919	0.917	-0.0870	-1.30
T3	117.46876	116.72184	1.006	0.0064	0.10
T4	188.15693	172.24677	1.092	0.0883	1.33
T5	171.04193	156.70796	1.091	0.0875	1.31
Sum	.	.	.	.	1.11

Table 11. Likelihood ratio test (df=3) decomposition of eigenvalues of correlation matrix for the pasture forage data fitted with a two-factor factor analytic (FA(2)) model.

Dimension	E(Fitted)	E(Full)	ER	Log(ER)	LR
D1	2.5959044	2.4523415	1.059	0.0569	0.85
D2	1.3151454	1.5197386	0.865	-0.1446	-2.17
D3	0.8912489	0.8462221	1.053	0.0518	0.78
D4	0.1935569	0.1776492	1.090	0.0858	1.29
D5	0.0041443	0.0040487	1.024	0.0234	0.35
Sum	.	.	.	.	1.10

Figure 1. Endpoints of 6 time vectors plotted in the space of the first two eigenvectors and then connected in sequential order for a first-order autoregressive process.
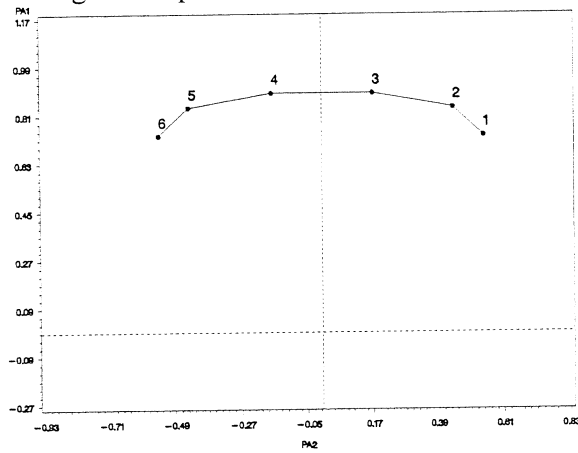
Figure 2. Endpoints of 6 time vectors plotted in the space of the last two eigenvectors and then connected in sequential order for a first-order autoregressive process.
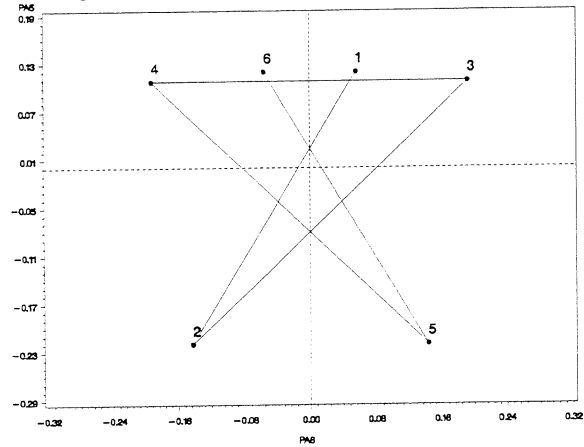

Figure 3. Endpoints of the 6 time vectors for the ARH(1) data fit with a homogeneous compound symmetry (CS) model plotted in the space of the first two principal axes.
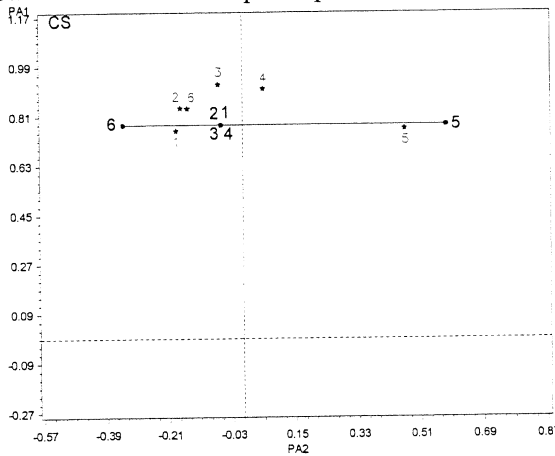
Figure 4. Endpoints of the 6 time vectors for the ARH(1) data fit with a heterogeneous first-order autoregressive (ARH(1)) model plotted in the space of the first two principal axes.
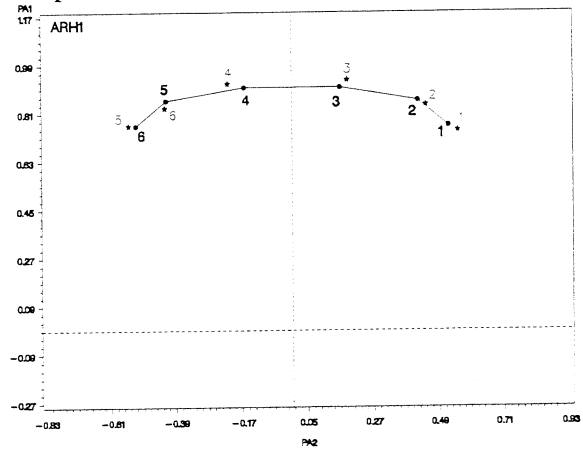
Figure 5. Endpoints of the 6 time vectors for the pasture forage data fit with a heterogeneous compound symmetry (CSH) model plotted in the space of the first two principal axes.
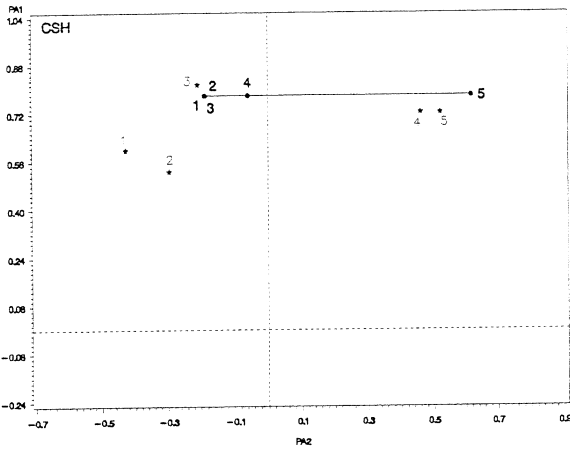

Figure 6. Endpoints of the 6 time vectors for the pasture forage data fit with a heterogeneous first-order autoregressive (ARH(1)) model plotted in the space of the first two principal axes.
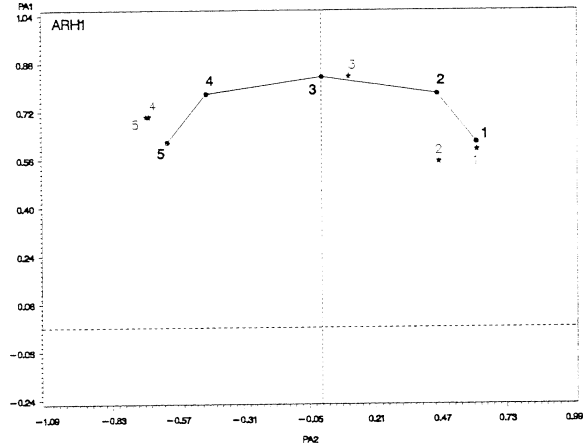

Figure 7. Endpoints of the 6 time vectors for the pasture forage data fit with a heterogeneous first-order autoregressive (ARH(1)) model plotted in the space of the last two principal axes.
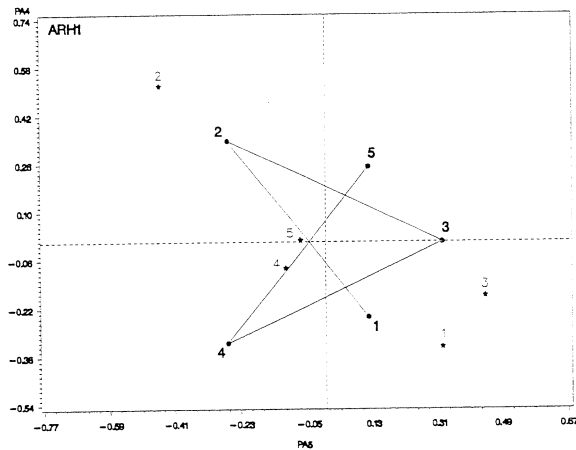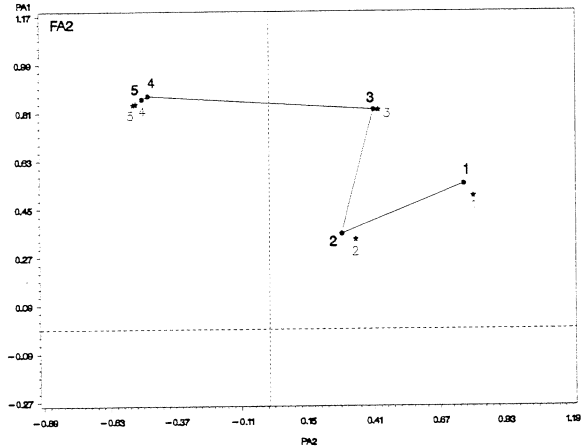

Figure 8. Endpoints of the 6 time vectors for the pasture forage data fit with a two-factor factor analytic (FA(2)) model plotted in the space of the first two principal axes.

Figure 9. Endpoints of the 6 time vectors for the pasture forage data fit with a two-factor factor analytic (FA(2)) model plotted in the space of the last two principal axes.