

This article was downloaded by: [Pioneer Hi-Bred International]

On: 10 April 2015, At: 14:29

Publisher: Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK

## The American Statistician

Publication details, including instructions for authors and subscription information:

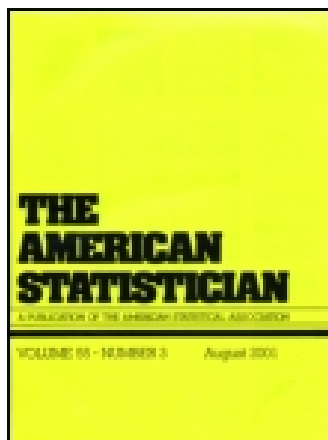
<http://amstat.tandfonline.com/loi/utas20>

### Viewing the Correlation Structure of Longitudinal Data through a PRISM

Dale L. Zimmerman<sup>a</sup>

<sup>a</sup> Department of Statistics and Actuarial Science, University of Iowa, Iowa City, IA, 52242

Published online: 17 Feb 2012.



To cite this article: Dale L. Zimmerman (2000) Viewing the Correlation Structure of Longitudinal Data through a PRISM, The American Statistician, 54:4, 310-318

To link to this article: <http://dx.doi.org/10.1080/00031305.2000.10474565>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://amstat.tandfonline.com/page/terms-and-conditions>

# Statistical Computing and Graphics

## Viewing the Correlation Structure of Longitudinal Data Through a PRISM

Dale L. ZIMMERMAN

Analysis of longitudinal data using a general linear mixed model requires the specification of a form for the covariance matrix of within-subject observations. Graphical diagnostics, such as the scatterplot matrix, can be of substantial help in making this specification. I introduce another graphical diagnostic, the Partial-Regression-on-Intervenors Scatterplot Matrix (PRISM), which complements the ordinary scatterplot matrix and which is more useful for identifying certain kinds of correlation structures. PRISMs corresponding to several commonly used correlation structures are displayed. The PRISM's usefulness in model specification is illustrated with an example of longitudinal data from a 100-kilometer road race.

**KEY WORDS:** Antedependence; Autoregressive model; Covariance matrix; Draftsman's display; Graphical diagnostic; Mixed model; PROC MIXED; Random coefficients; Scatterplot matrix.

### 1. INTRODUCTION

Consider a situation in which repeated measurements of a continuous response variable are taken over time on each of several subjects. Such *longitudinal data* can be analyzed statistically in a variety of ways; see Diggle, Liang, and Zeger (1994) for a comprehensive treatment. In the past 20 years or so, an approach for analyzing longitudinal data that is based on mixed linear models has become quite popular and has been implemented in major statistical software packages; for example, PROC MIXED of SAS (SAS Institute Inc. 1996). Application of a mixed linear model to a particular longitudinal dataset requires the specification of the model's mean structure and covariance structure. Often, the study's objectives largely dictate the choice of mean structure; for example, if the goal of the study is to see how the growth of animals is affected over time by different treatments, then the initial specification of mean structure would probably include treatment effects, one or more time effects (e.g., linear and quadratic terms), and effects for time-by-treatment interaction. Graphical techniques, such as a plot

of subject profiles over time, can be used to inform or confirm this choice; these plots have additional capabilities such as revealing anomalous observations.

However, an appropriate covariance structure is often more difficult to specify than a mean structure. Typically, the study's objectives do not dictate this choice; nevertheless, it is important because an appropriately parsimonious choice of covariance structure can substantially improve the efficiency of inferences made about the mean structure and provide better estimates of standard errors of estimated mean parameters (Diggle et al. 1994). It is often reasonable to assume that responses from different subjects are independent and that the within-subject covariance structure is homogeneous across subjects, or at least is homogeneous across subjects within specified subgroups (e.g., treatment groups). However, it typically is not at all clear, a priori, how this common within-subject covariance structure should be specified. The problem is not a shortage of possible structures. PROC MIXED, for example, allows the user to choose from no less than 20 distinct covariance structures (e.g., compound symmetry, first-order autoregressive, Toeplitz, and Huynh-Feldt), and additional possibilities were described by Jennrich and Schluchter (1986); Muñoz, Carey, Schouten, Segal, and Rosner (1992); and Núñez-Antón and Zimmerman (2000). Rather, the problem is to reduce the overall number of covariance structures that need to be fit to the data and, subsequently, to assess whether the structure judged to fit "best" (according, typically, to some numerical criterion) actually provides a reasonable fit. An examination of the sample covariance and sample correlation matrices (or, if the data are unbalanced, the matrix of sample variances and correlations computed using all available data) may prove helpful in this regard, but graphical techniques are likely to be even more useful. In addition to reducing the number of covariance structures that need to be considered, a graphical technique may reveal the existence of subjects exhibiting dependence markedly different from the rest or indicate other interesting data attributes.

A covariance structure can be decomposed into two components: a variance structure and a correlation structure. Many diagnostics that are useful for specifying mean structure can also convey useful information about variances' behavior over time; a series of boxplots of the marginal distributions of responses at time  $t_i$ , plotted against  $i$ , is an example. Here we focus on diagnostics for the correlation structure only. Diggle et al. (1994) and Dawson, Gennings,

Dale L. Zimmerman is Professor, Department of Statistics and Actuarial Science, University of Iowa, Iowa City, IA 52242 (E-mail: dzimmer@stat.uiowa.edu). This work was partially supported by National Science Foundation grant 9628612.

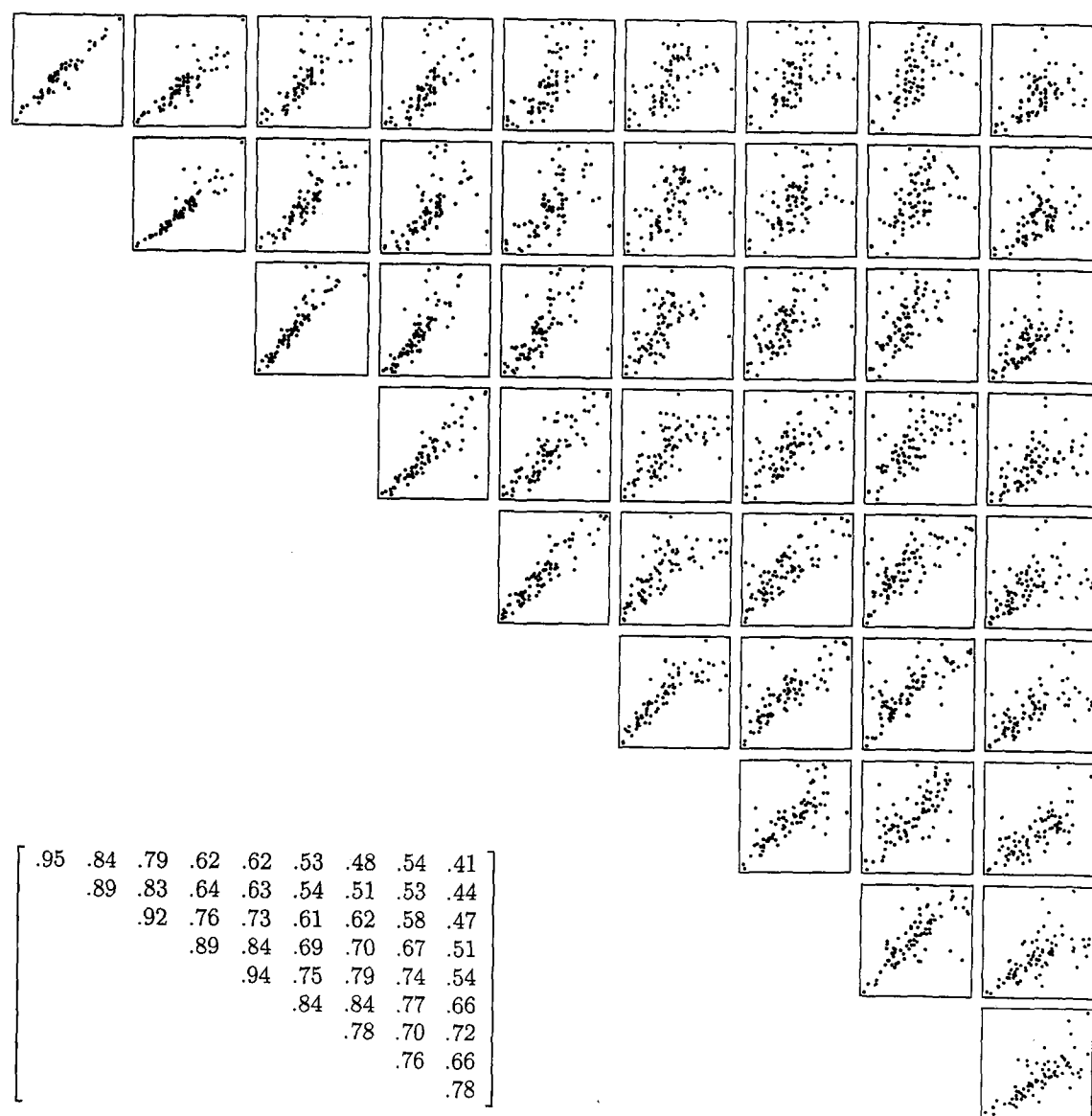


Figure 1. Ordinary Scatterplot Matrix and Sample Correlation Matrix for the Standardized 100-km Race Data.

and Carter (1997) reviewed several methods for diagnosing correlation structure. The most widely used method is the *scatterplot matrix* or *draftsman's display*, which is a two-dimensional array of pairwise scatterplots of standardized responses (or of certain derived quantities such as residuals). Figure 1 displays a scatterplot matrix for the (standardized) 100-kilometer race data discussed in Section 4. The scatterplot matrix is a graphical equivalent of the sample correlation matrix and is most useful for diagnosing correlation structure when measurement times are common across subjects and equally spaced, though mild departures from these ideals can be tolerated.

Although the scatterplot matrix is generally useful for diagnosing correlation structure, alternative procedures can be complementary or even more informative in some cases. This article introduces a new graphical diagnostic called the Partial Regression-on-Intervenors Scatterplot Matrix (PRISM), which is particularly adept at identifying autoregressive and certain other types of serial correlation struc-

tures. The acronym PRISM has some semantic substance: as a prism separates visible light into its components, so a PRISM can separate the dependence structure among within-subject responses into components that are easier to understand.

Section 2 briefly reviews the general mixed linear model approach to longitudinal data analysis and describes some commonly used correlation structures. Section 3 describes the PRISM in detail, and its distinctive features are illustrated for several correlation structures. Section 4 further illustrates the PRISM's use through the analysis of longitudinal data from a 100-kilometer race. Concluding remarks are made in Section 5.

## 2. THE MIXED MODEL APPROACH

Consider a longitudinal study with  $m$  subjects. Let  $\mathbf{y}_i = (y_{i1}, \dots, y_{in_i})'$  be a response vector of  $n_i$  measurements on the  $i$ th subject, and let  $\mathbf{t}_i = (t_{i1}, \dots, t_{in_i})'$  be the corresponding vector of measurement times. Suppose that we

Table 1. Six Common Correlation Structures

Correlation structure	Parameter space	
1. First-order autoregressive [AR(1)]	$\rho_{jk}$	$= \rho^{t_k - t_j}, 0 < \rho < 1$
2. Second-order autoregressive [AR(2)]	$\rho_{jk}$	$= [1 + (1 + \phi_2)(1 - \phi_2)^{-1}(k - j)](\phi_1/2)^{k-j}$ if $\phi_1^2 + 4\phi_2 = 0$ $= \frac{(1 - \gamma_2^2)\gamma_1^{k-j+1} - (1 - \gamma_1^2)\gamma_2^{k-j+1}}{(\gamma_1 - \gamma_2)(1 + \gamma_1\gamma_2)}$ otherwise where $\gamma_1 = \frac{-2\phi_2}{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}$ and $\gamma_2 = \frac{-2\phi_2}{\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}$ , $\phi_1 + \phi_2 < 1$ , $\phi_2 - \phi_1 < 1$ , $ \phi_2  < 1$
3. First-order antedependence [AD(1)]	$\rho_{jk}$	$= \prod_{l=j}^{k-1} \rho_l, \quad -1 < \rho_l < 1 \quad \forall l$
4. First-order moving average [MA(1)]	$\rho_{jk}$	$= \rho$ if $k - j = 1$ $= 0$ otherwise, where $-.5 < \rho < .5$
5. Compound symmetry (CS)	$\rho_{jk}$	$= \rho, \quad -\frac{1}{n-1} < \rho < 1$
6. Linear random coefficients (LRC)	$\rho_{jk}$	$= \frac{\sigma_{00} + \sigma_{01}(t_j + t_k) + \sigma_{11}t_j t_k}{\sqrt{\sigma^2 + \sigma_{00} + 2\sigma_{01}t_j + \sigma_{11}t_j^2} \sqrt{\sigma^2 + \sigma_{00} + 2\sigma_{01}t_k + \sigma_{11}t_k^2}}, \quad \sigma_{00} > 0, \sigma_{11} > 0, \sigma_{00}\sigma_{11} - \sigma_{01}^2 > 0$

NOTE: For these representations we assume that measurement times are common across subjects—that is,  $\mathbf{t}_i \equiv \mathbf{t} = (t_j)$  and  $n_i \equiv n$ . Moreover, for the AR(2) and MA(1) structures we assume equally spaced observations—that is,  $t_{j+1} - t_j = 1 \quad \forall j$ . Correlations are denoted as  $\rho_{jk} = \text{corr}(y_{ij}, y_{ik}), j < k$ . Model parameters are denoted by Greek symbols. Parameter spaces that ensure positive definiteness are given with each structure.

also observe a  $p$ -vector of covariates,  $\mathbf{x}_{ij}$ , associated with  $y_{ij}$ , and put  $\mathbf{X}_i = (\mathbf{x}_{i1}, \dots, \mathbf{x}_{in_i})'$ . A general mixed linear model (Harville 1977; Laird and Ware 1982) for the responses is given by

$$\mathbf{y}_i \sim \text{independent } N(\mathbf{X}_i\boldsymbol{\beta}, \boldsymbol{\Sigma}_i(\boldsymbol{\theta})), \quad (1)$$

where  $\boldsymbol{\beta}$  is a  $p$ -vector of fixed, unknown, and typically unrestricted parameters;  $\boldsymbol{\Sigma}_i(\boldsymbol{\theta})$  is a  $n_i \times n_i$  covariance matrix; and  $\boldsymbol{\theta}$  is a  $q$ -vector of unknown parameters, restricted to a parameter space  $\Theta$  that is either the set of all  $\boldsymbol{\theta}$ -vectors for which all  $\boldsymbol{\Sigma}_i$  are positive definite or some subset of that set. Note that the  $\boldsymbol{\Sigma}_i$ 's are assumed to be of the same basic form—that is, the same parametric covariance function determines the elements of each—but generally they may not be equal nor of the same dimensions; an exception occurs when measurement times are common across subjects.

Estimation of the parameters of model (1) is typically carried out by the methods of maximum likelihood or residual maximum likelihood. As noted previously, these methods are implemented in PROC MIXED for a variety of covariance structures. Table 1 shows six frequently used correlation structures. The first-order and second-order autoregressive [AR(1) and AR(2)] structures originate from stationary time series analysis. The first-order antedependence [AD(1)] structure, introduced by Gabriel (1962), is generally nonstationary, as it allows variances to be time-dependent and correlations between pairs of observations equidistant in time to be unequal. More specifically, it is characterized by correlations on the third and higher diagonals of the correlation matrix that are equal to certain products of the correlations on the second diagonal. Clearly the AR(1) is a special case. The AD(1) structure is the most general case of first-order Markov structure, which means that responses at any two times are conditionally independent given a response at any intervening time. The AD(1) structure is itself a special case of the  $k$ th-order antedependence [AD( $k$ )], or  $k$ th-order Markov structure, which is characterized by responses that are conditionally independent

given the responses at any  $k$  intervening times. We do not include the AD( $k$ ) in Table 1 because the corresponding correlations cannot be expressed in a simple form for  $k > 1$ . However, the correlation structure can be deduced from the AD( $k$ ) model equations

$$\begin{aligned} y_1 &= \mu_1 + \epsilon_1, \\ y_t &= \mu_t + \sum_{l=1}^{k^*} \phi_{lt}(y_{t-l} - \mu_{t-l}) + \epsilon_t \quad (t = 2, \dots, n) \end{aligned} \quad (2)$$

(Zimmerman and Núñez-Antón 1997). Here  $y_t$  is a generic subject's response at time  $t$ ,  $\mu_t = E(y_t)$ ,  $k^* = \min(k, t-1)$ , the  $\epsilon_t$ 's are independent normal random variables with zero means and possibly time-dependent variances, and the  $\phi_{lt}$ 's are arbitrary constants. An even more general Markovian structure than the AD( $k$ ) is variable-order antedependence (VAD, Macchiavelli and Arnold 1994), in which the number of intervening responses that must be conditioned on to achieve independence depends on the measurement time (or equivalently,  $k^*$  in (2) is a function of  $t$ ).

The remaining three structures listed in Table 1 are non-Markovian. The first-order moving average [MA(1)] structure originates from stationary time series analysis. The compound symmetry (CS) and linear random coefficients (LRC) structures arise from supposing that  $\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \mathbf{e}_i$ , where  $\mathbf{Z}_i = \mathbf{1}_{n_i}$  for CS and  $\mathbf{Z}_i = (\mathbf{1}_{n_i}, \mathbf{t}_i)$  for LRC,

$$\mathbf{b}_1, \dots, \mathbf{b}_n \text{ are iid } N_2\left(\mathbf{0}, \begin{pmatrix} \sigma_{00} & \sigma_{01} \\ \sigma_{01} & \sigma_{11} \end{pmatrix}\right),$$

$\mathbf{e}_1, \dots, \mathbf{e}_n$  are independently distributed as  $N_{n_i}(\mathbf{0}, \sigma^2\mathbf{I})$ , and  $\mathbf{b}_i$  and  $\mathbf{e}_j$  are independent for all  $i$  and  $j$ .

### 3. THE PRISM

Before constructing the PRISM it is advisable to standardize the response variables to prevent possible differences in the variances of responses over time from obfuscating the correlation structure. For simplicity of pre-

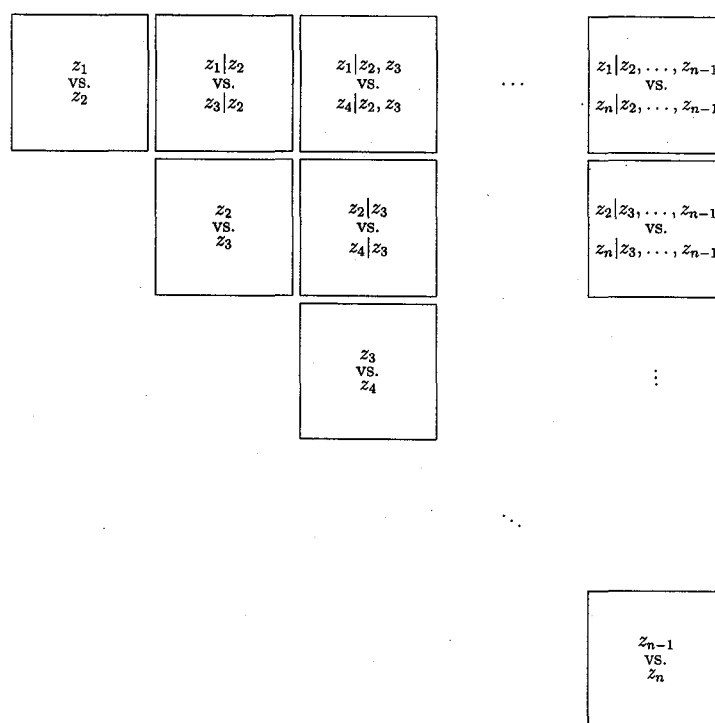


Figure 2. Layout of Partial Regression Plots in a PRISM.

sensation, let us now assume that there are no treatments or other observed covariates, and that measurement times  $t_1, t_2, \dots, t_n$  (not necessarily equally spaced) are common across subjects, save possibly for a relatively small proportion of missing data. Let  $\bar{y}_{\cdot j}$  and  $s_{\cdot j}$  denote the sample mean and sample standard deviation, respectively, of the nonmissing responses at time  $j$ . Define the standardized responses

$$z_{ij} = \frac{y_{ij} - \bar{y}_{\cdot j}}{s_{\cdot j}}$$

for  $i = 1, \dots, m$  and  $j = 1, \dots, n$ . Dawson et al. (1997) showed that the sample correlation structure of the standardized observations is identical to that of the original observations. Note, for reasons that will become apparent in the next paragraph, that the same type of identity holds for the sample partial correlation structure also.

The PRISM is constructed as the upper right "wedge" of a rectangular array of certain partial regression (or added variable) plots, laid out as indicated in Figure 2. The main diagonal of plots are ordinary scatterplots of  $z_j$  (the collection of nonmissing standardized responses at time  $j$ ) against  $z_{j+1}$  ( $j = 1, \dots, n-1$ ). To provide for a precise description of the remaining plots, define  $\hat{\alpha}_j^{(+k)}$  as the least squares estimate of the intercept, and  $\hat{\beta}_{j,j+l}^{(+k)}$  as the least squares estimate of the slope coefficient corresponding to  $z_l$  ( $l = 1, \dots, k$ ), in the regression of  $z_j$  on  $z_{j+1}, \dots, z_{j+k}$  ( $j = 1, \dots, n-k; k = 1, \dots, n-1$ ). Similarly, define  $\hat{\alpha}_j^{(-k)}$  as the least squares estimate of the intercept, and  $\hat{\beta}_{j,j-l}^{(-k)}$  as the least squares estimate of the slope coefficient corresponding to  $z_{j-l}$  ( $l = 1, \dots, k$ ), in the regression of  $z_j$  on  $z_{j-1}, \dots, z_{j-k}$  ( $j = k+1, \dots, n; k = 1, \dots, n-1$ ). Then, the second diagonal of plots in the PRISM are partial re-

gression plots of standardized responses lagged two times apart, adjusted for the standardized response at the intervening time; that is, plots of  $z_j - \hat{\alpha}_j^{(+1)} - \hat{\beta}_{j,j+1}^{(+1)} z_{j+1}$  versus  $z_{j+2} - \hat{\alpha}_{j+2}^{(-1)} - \hat{\beta}_{j+2,j+1}^{(-1)} z_{j+1}$  ( $j = 1, \dots, n-2$ ). The third diagonal of plots are partial regression plots of standardized responses lagged three times apart, adjusted for the standardized responses at the two intervening times; that is, plots of  $z_j - \hat{\alpha}_j^{(+2)} - \hat{\beta}_{j,j+1}^{(+2)} z_{j+1} - \hat{\beta}_{j,j+2}^{(+2)} z_{j+2}$  versus  $z_{j+3} - \hat{\alpha}_{j+3}^{(-2)} - \hat{\beta}_{j+3,j+2}^{(-2)} z_{j+2} - \hat{\beta}_{j+3,j+1}^{(-2)} z_{j+1}$  ( $j = 1, \dots, n-3$ ). In general, the plot in row  $j$  and column  $k$  ( $k \geq j$ ) is the partial regression plot of standardized response variables  $z_j$  and  $z_{k+1}$  adjusted for standardized responses at the intervening times  $t_{j+1}, t_{j+2}, \dots, t_k$ . Thus, the PRISM is the graphical equivalent of a matrix of certain partial correlations; specifically, the  $(j, k)$ th plot in the array displays points whose ordinary correlation is the partial correlation between  $z_j$  and  $z_{k+1}$  adjusted for all standardized responses at intervening times  $t_{j+1}, t_{j+2}, \dots, t_k$ . Random scatter in the  $(j, k)$ th plot indicates that  $z_j$  and  $z_{k+1}$  are conditionally independent, given the intervening responses, whereas departures from random scatter indicate conditional dependence.

Several commonly used correlation structures have distinctive associated partial correlation structures, and it is these structures that the PRISM is particularly effective at identifying. Figures 3 and 4 show the ordinary scatterplot matrix (OSM) and PRISM for simulated longitudinal data having several correlation structures. In each case, the data are a random sample of size  $m = 100$  from a six-dimensional normal distribution with mean vector  $\mathbf{0}$  and unit variances, and the measurement times are  $t_i = i$  ( $i = 1, \dots, 6$ ). These examples highlight the relative strengths and weaknesses of the OSM and PRISM and indi-

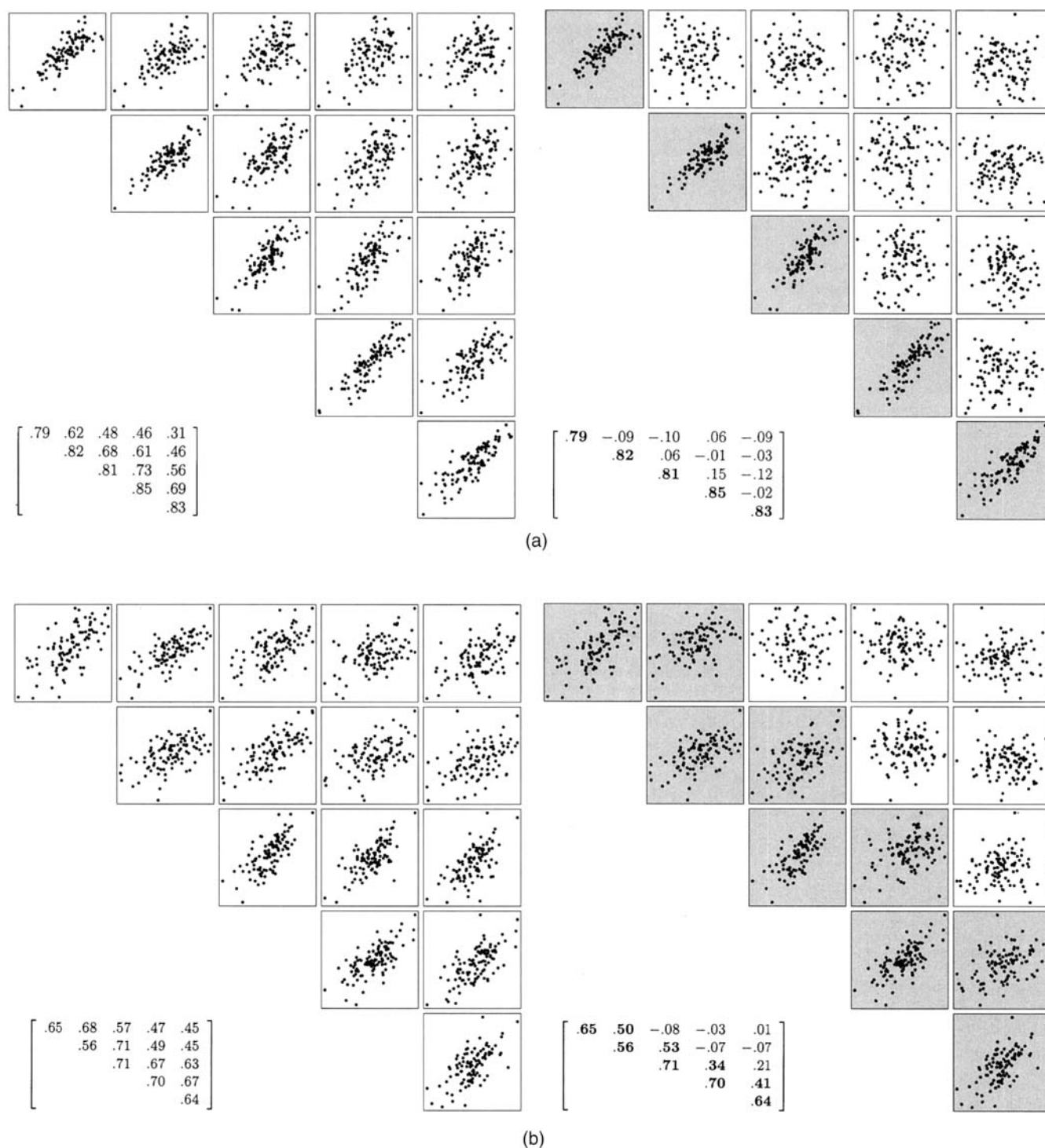


Figure 3. Ordinary Scatterplot Matrices (left) and PRISMs (right), and Corresponding Sample Correlation and Partial Correlation Matrices, for Simulated Data Having Selected Correlation Structures: (a) AR(1) with  $\rho = .8$  and (b) AR(2) with  $\phi_1 = .3$ ;  $\phi_2 = .5$ . Those partial correlations on intervenors significantly different from zero at a Bonferroni-adjusted 5% level appear in bold type, and the corresponding plots in the PRISMs are shaded.

cate how they can complement one another for the purpose of structure identification.

The OSMs for the AR(1) and AR(2) structures (Figures 3a and 3b) are qualitatively very similar. Both show the persistence, over time, of the correlations, which attenuate but do not vanish as one moves away from the main diagonal. Both also reveal a constancy of correlation strength within

diagonals, which is a manifestation of the stationarity of these two models. The corresponding PRISMs, however, differ markedly with respect to their second diagonals. Plots in the second diagonal of the AR(1)'s PRISM exhibit random scatter, but as a consequence of the AR(2)'s nonzero partial correlations of the form  $\text{corr}(z_j | z_{j+1}, z_{j+2} | z_{j+1})$ , their counterparts in the AR(2)'s PRISM do not. The third

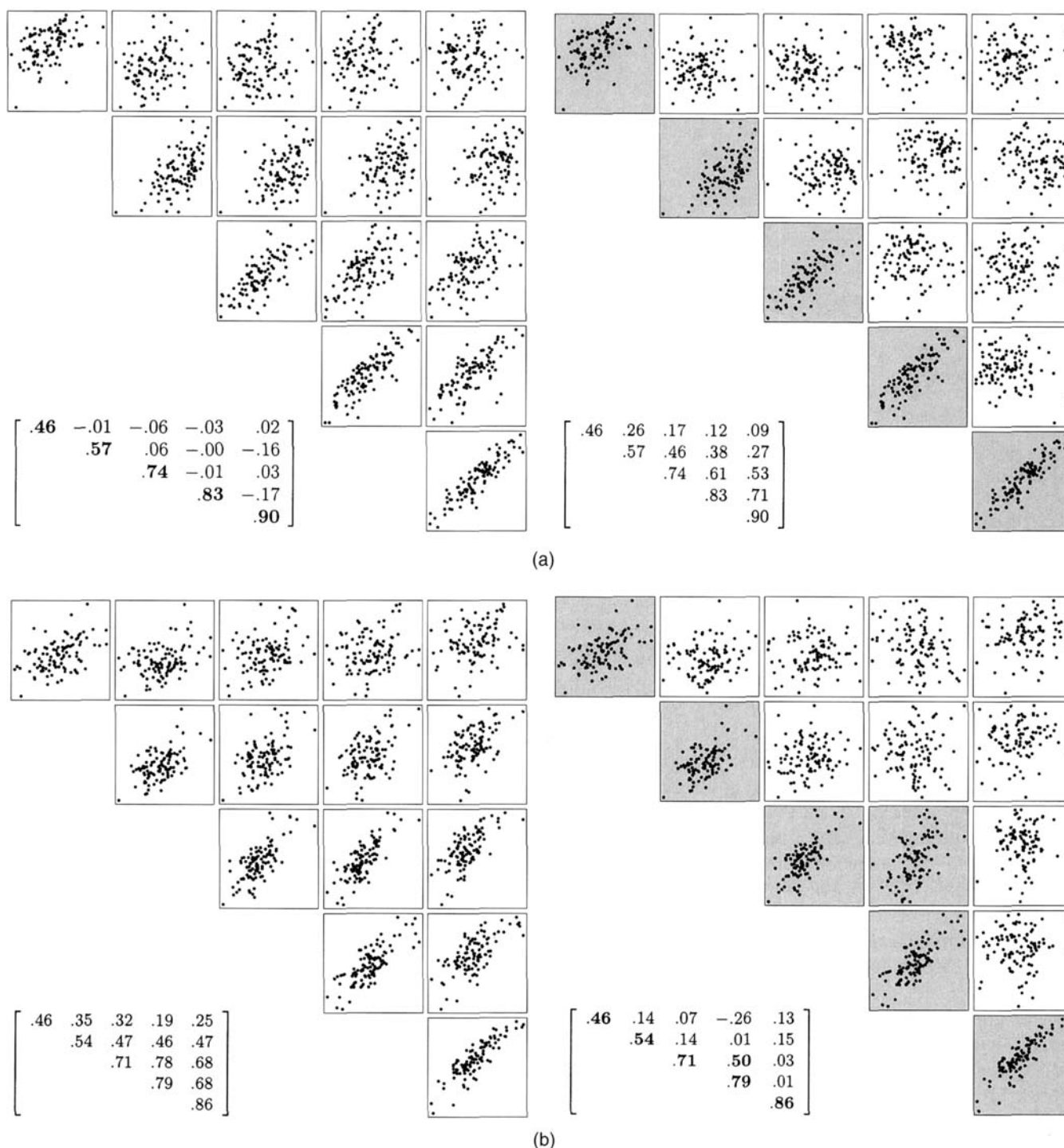


Figure 4. Ordinary Scatterplot Matrices (left) and PRISMs (right), and Corresponding Sample Correlation and Partial Correlation Matrices, for Simulated Data Having Selected Correlation Structures: (a) AD(1) with  $(\rho_1, \rho_2, \rho_3, \rho_4, \rho_5) = (.5, .6, .7, .8, .9)$ ; (b) VAD of order 1 for times 2, 3, 4, and 6, and order 2 for time 5, with  $(\phi_{12}, \phi_{13}, \phi_{14}, \phi_{15}, \phi_{16}, \phi_{25}) = (.5, .6, .7, .8, .9, .7)$ . Those partial correlations on intervenors significantly different from zero at a Bonferroni-adjusted 5% level appear in bold type, and the corresponding plots in the PRISMs are shaded.

and higher diagonals of both PRISMs exhibit random scatter, however.

The OSMs and PRISMs corresponding to two more general Markovian dependence structures, AD(1) and VAD, are given in Figures 4a and 4b. The two OSMs are very similar. Both show the persistence of correlation as one moves away from the main diagonal and, in contrast to the pre-

vious two cases, an increase in correlation strength as one moves down any particular diagonal. The PRISMs also are quite similar to one another, but they differ in the plot in row 3, column 4. This plot exhibits random scatter for the AD(1) but not for the VAD, reflecting the latter's nonzero value of  $\text{corr}(z_3|z_4, z_5|z_4)$ .

The previous examples demonstrate clearly the ability of the PRISM to identify and distinguish between autoregres-

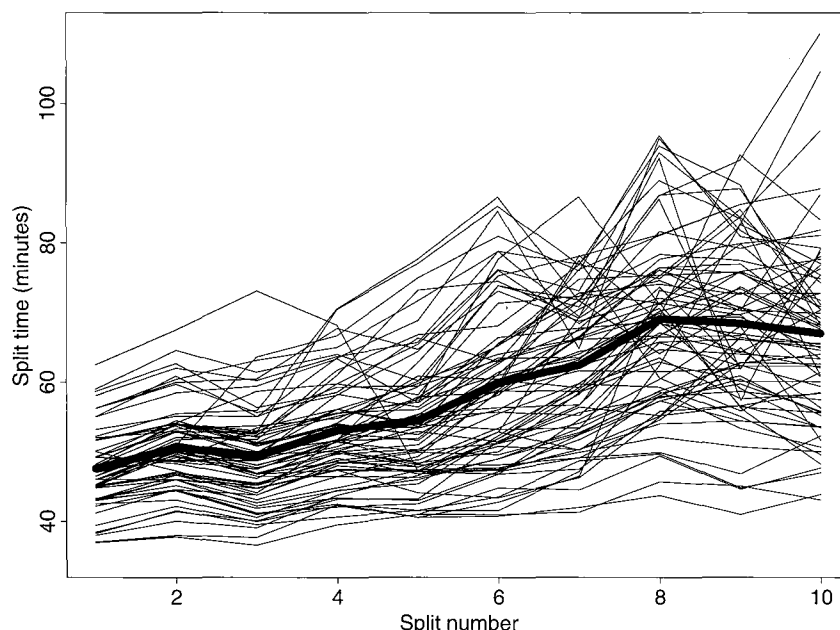


Figure 5. Profile Plot of 100-km Race Data. The thick line indicates the mean split time for each 10-km section.

sive structures and other Markovian dependence structures. For non-Markovian dependence, however, the OSM may be more informative than the PRISM. For example, the MA(1) and CS structures are more easily diagnosed from the OSM than from the PRISM. Moreover, neither the OSM nor the PRISM is particularly adept at identifying random coefficients models such as LRC.

These examples suggest that a better diagnosis of correlation structure occurs when the OSM and PRISM are used in tandem than when only one is used. Using the PRISM in addition to the OSM for specifying the correlation structure of longitudinal data would parallel a similar practice in time series analysis. There, good statistical practice includes an examination of both the sample autocovariance function (ACF) and sample partial autocovariance function (PACF). In fact, in the case of equally spaced longitudinal data from a stationary process, the superposition of plots of the OSM along diagonals yields  $n - 1$  scatterplots whose correlations coincide with those given by the ACF, and similarly superimposing plots of the PRISM yields  $n - 1$  scatterplots whose correlations coincide with those given by the PACF.

#### 4. EXAMPLE

Everitt (1994a, 1994b) reported some exploratory and graphical analyses of data from a 100-kilometer (km) road race held in 1984 in the United Kingdom. The data consist of the “split” times for 80 runners in each 10-km section of the race. A profile plot of the split times (Figure 5) indicates that: (a) the mean tends to increase over the first 80 km of the race but then levels off; (b) the variance tends to increase as the race progresses; and (c) the behavior of many of the runners is more erratic, in the sense that consecutive same-runner split times fluctuate more, in the later sections of the race. The OSM of the standardized responses and the sample correlation matrix (Figure 1) show that: (a) correlations are positive and quite large; (b) the correlation between the

split time for any fixed 10-km section and split times for successive sections tends to decrease monotonically; and (c) correlations between split times of consecutive sections are not as large near the end of the race as they were earlier. This last feature is a consequence of the increased erraticity near the end of the race noted previously.

Figure 6 displays a PRISM of the same data and the corresponding matrix of partial correlations on intervenors. Apart from plots on the main diagonal, the plot in row 6, column 7, and possibly a few others, the plots appear as random scatter. This can be investigated more formally via standard size-.05 likelihood ratio tests for the corresponding partial correlation coefficients. Using a Bonferroni adjustment, the only partial correlations on intervenors significantly different from zero are  $\text{corr}(z_8|z_7, z_6|z_7)$  and  $\text{corr}(z_{10}|(z_8, z_9), z_7|(z_8, z_9))$ . Nearly significant are  $\text{corr}(z_3|z_4, z_5|z_4)$  and  $\text{corr}(z_{10}|(z_6, z_7, z_8, z_9), z_5|(z_6, z_7, z_8, z_9))$ , but the PRISM suggests that the near-significance of the first of these is due to one aberrant observation.

The graphical analysis suggests that antedependence models up to order two should certainly be fit to the data, and antedependence models of orders three, four, and five should probably be fit as well. Variable-order antedependence models could also be considered but we do not do so here. For illustration we fit antedependence models of order up to five as well as several covariance models available in PROC MIXED. As noted previously, the sample variances are not constant, hence we only consider models that allow for heterogeneous variances. Fits of covariance models are compared using two widely used information criteria in larger-is-better form:  $\text{AIC} = l_R(\hat{\theta}) - q$  and  $\text{BIC} = l_R(\hat{\theta}) - \frac{q}{2} \log(N - p)$ . Here  $l_R$  is the residual log-likelihood, and  $\hat{\theta}$  and  $q$  are the residual maximum likelihood (REML) estimate and dimension, respectively, of the vector of covariance parameters;  $N = 800$  is the total sample size; and  $p = 10$  is the number of measurement times.



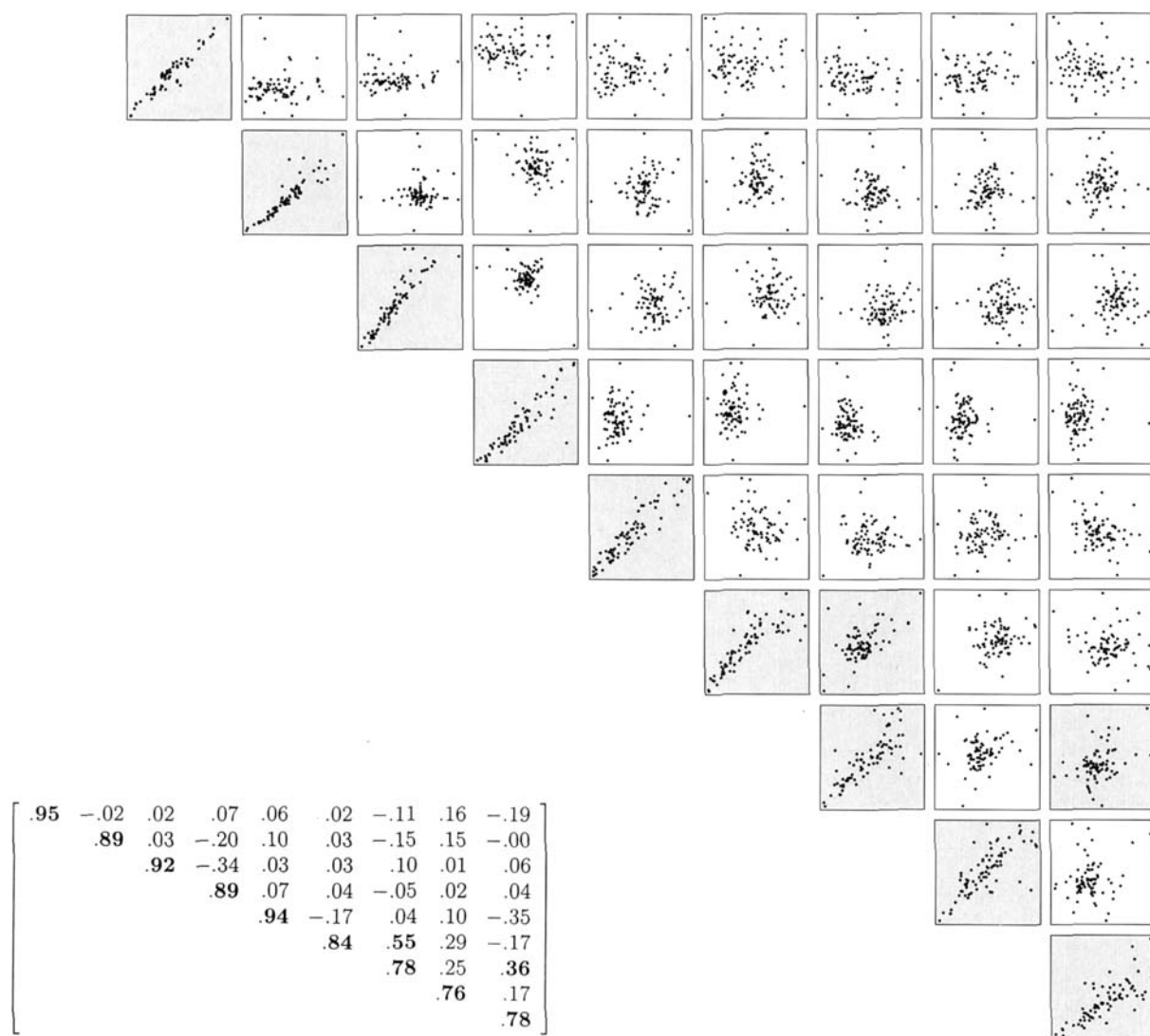


Figure 6. PRISM of 100-km Race Data and Corresponding Sample Partial Correlation Matrix. Those partial correlations on intervenors significantly different from zero at a Bonferroni-adjusted 5% level appear in bold type, and the corresponding plots in the PRISM are shaded.

Another criterion used to compare fits of covariance models is a residual likelihood ratio test (RLRT). A RLRT is conducted by differencing the minimized values of  $-2l_R$  for two nested models and comparing this value to the  $\chi^2$ -distribution with degrees of freedom equal to the difference in the number of covariance parameters.

Table 2 gives  $q$ , AIC, BIC, and  $-2l_R$  for each fitted model. BIC, which penalizes models with more parameters more severely, supports the more parsimonious ARH(1) and AD(1) models. On the basis of both AIC and the RLRT, however, the AD(3) model is superior. The AD(2) model is rejected in favor of the AD(3) by the RLRT, and the AD(3)

Table 2. REML Information Criteria and Likelihood Ratio Tests of Covariance Structures for the 100-Kilometer Race Data

Structure	$q$	AIC	BIC	$-2l_R$	CM	$\nu$	$\chi^2$	$P$
CSH	11	-2558.6	-2584.3	5095.2				
LRC	4	-2566.6	-2575.9	5125.1				
QRC	7	-2525.0	-2541.4	5036.1	LRC	3	89.0	0
ARH(1)	11	-2395.7	-2421.4	4769.4				
AD(1)	19	-2377.0	-2421.4	4716.0	ARH(1)	8	53.4	0
AD(2)	27	-2361.1	-2424.2	4668.2	AD(1)	8	47.8	0
AD(3)	34	-2357.3	-2436.7	4646.6	AD(2)	7	21.6	0
AD(4)	40	-2361.0	-2454.4	4642.0	AD(3)	6	4.6	.60
AD(5)	45	-2361.1	-2466.2	4632.1	AD(3)	11	14.5	.21
UN	55	-2365.2	-2493.7	4620.4	AD(3)	21	26.2	.20

NOTE: Here,  $q$  is the number of covariance parameters, CM is the comparison model,  $\nu$  is the degrees of freedom,  $\chi^2$  is the test statistic, and  $P$  is the  $p$  value for the residual likelihood ratio test. Also, CSH is a compound symmetry correlation structure with heterogeneous variances, LRC and QRC are the linear and quadratic random coefficients models, respectively, ARH(1) is an AR(1) correlation structure with heterogeneous variances, and UN is a completely unstructured covariance model.

is not rejected when tested against either the AD(4), AD(5), or UN models.

In this example the graphical analysis comports well with conclusions based on AIC and the RLRT. It also points to some possible further analyses, such as investigating the outlier evident in the fourth row and fourth column of plots in the PRISM, or searching for a substantive explanation for the positive, significantly-different-from-zero partial correlations on intervenors in the later sections of the race.

## 5. DISCUSSION

We have introduced a new graphical diagnostic, the PRISM, which complements the ordinary scatterplot matrix and other existing methods for diagnosing the correlation structure of longitudinal data. Neither scatterplot matrix is uniformly superior to the other; the PRISM is better at diagnosing autoregressive and more general Markov-type structures while the ordinary scatterplot matrix is better for other structures. Consequently, I recommend that the ordinary scatterplot matrix and PRISM be used in tandem, much as plots of the sample autocovariance function and sample partial autocovariance function are used in tandem for model identification in time series analysis. By heeding this advice, some inappropriate models for the covariance structure may be eliminated from consideration, thereby reducing the number the analyst needs to fit. Moreover, these plots may permit an assessment of whether the model(s) that minimizes a numerical goodness-of-fit criterion actually provide(s) a reasonable fit. The graphical analysis could be made even more penetrating by dynamically linking the plots within the PRISM to one another and by dynamically linking the the OSM and PRISM to each other and to the profile plot. Also, because the OSM and the PRISM share the same main diagonal, and each (by virtue of the symmetry of a correlation matrix about its main diagonal) utilizes only half of the remaining positions in the scatterplot ma-

trix, the two could be combined into a single scatterplot matrix, if desired.

[Received June 1999. Revised April 2000.]

## REFERENCES

- Dawson, K. S., Gennings, C., and Carter, W. H. (1997), "Two Graphical Techniques Useful in Detecting Correlation Structure in Repeated Measures Data," *The American Statistician*, 51, 275-283.
- Diggle, P. J., Liang, K. Y., and Zeger, S. L. (1994), *Analysis of Longitudinal Data*, New York: Oxford University Press.
- Everitt, B. S. (1994a), "Exploring Multivariate Data Graphically: A Brief Review with Examples," *Journal of Applied Statistics*, 21, 63-94.
- (1994b), *A Handbook of Statistical Analyses Using S-Plus*, London: Chapman and Hall.
- Gabriel, K. R. (1962), "Ante-dependence Analysis of an Ordered Set of Variables," *Annals of Mathematical Statistics*, 33, 201-212.
- Harville, D. A. (1977), "Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems," *Journal of the American Statistical Association*, 72, 320-340.
- Jennrich, R. L., and Schluchter, M. D. (1986), "Unbalanced Repeated-Measures Models with Structured Covariance Matrices," *Biometrics*, 42, 805-820.
- Laird, N. M., and Ware, J. H. (1982), "Random-Effects Models for Longitudinal Data," *Biometrics*, 38, 963-974.
- Macchiavelli, R. E., and Arnold, S. F. (1994), "Variable Order Antedependence Models," *Communications in Statistics—Theory and Methods*, 23, 2683-2699.
- Muñoz, A., Carey, V., Schouten, J. P., Segal, M., and Rosner, B. (1992), "A Parametric Family of Correlation Structures for the Analysis of Longitudinal Data," *Biometrics*, 48, 733-742.
- Núñez-Antón, V., and Zimmerman, D. L. (2000), "Modelling Nonstationary Longitudinal Data," *Biometrics*, 56, 699-705.
- SAS Institute, Inc. (1996), *SAS/STAT Software: Changes and Enhancements through Release 6.11*, Cary, NC: SAS Institute Inc.
- Zimmerman, D. L., and Núñez-Antón, V. (1997), "Structured Antedependence Models for Longitudinal Data," in *Modelling Longitudinal and Spatially Correlated Data. Methods, Applications, and Future Directions*, eds. T. G. Gregoire, D. R. Brillinger, P. J. Diggle, E. Russek-Cohen, W. G. Warren, and R. Wolfinger, New York: Springer-Verlag, pp. 63-76.