Random Effects and AR(1) Errors

ERIC M. CHI and GREGORY C. REINSEL*

For longitudinal data on several individuals, linear models that contain both random effects across individuals and autocorrelation in the within-individual errors are studied. A score test for autocorrelation in the within-individual errors for the "conditional independence" random effects model is first developed. An explicit maximum likelihood estimation procedure using the scoring method for the model with random effects and (autoregressive) AR(1) errors is then derived. Empirical Bayes estimation of the random effects and prediction of future responses of an individual based on this random effects with AR(1) errors model are also considered. A numerical example is presented to illustrate these methods.

KEY WORDS: Autoregressive model; Empirical Bayes estimate; Maximum likelihood estimation; Prediction; Score test.

1. INTRODUCTION

To model longitudinal data collected over time for each member of a group of experimental units, one must recognize the correlation between serial observations on the same experimental unit. For balanced and complete longitudinal or "growth curve" data, Potthoff and Roy (1964), Rao (1965, 1967), and Grizzle and Allen (1969) examined through multivariate analysis techniques the generalized multivariate analysis of variance model

$$\mathbf{y}_k = \mathbf{X}\mathbf{B}\mathbf{a}_k + \mathbf{e}_k, \qquad k = 1, \ldots, N, \tag{1.1}$$

where $\mathbf{y}_k$ is the $p \times 1$ vector of observations on individual k, $\mathbf{X}$ is the $p \times q$ within individual design matrix, $\mathbf{B}$ is the $q \times r$ matrix of unknown parameters, $\mathbf{a}_k$ is the $r \times 1$ vector of non–time-varying covariates associated with individual k, and the $\mathbf{e}_k$ are $p \times 1$ errors assumed to be independently distributed as multivariate normal with mean $\mathbf{0}$ and general $p \times p$ covariance matrix Σ.

In practice, longitudinal data are often unbalanced or incomplete; that is, all individuals are not observed at the same number of time points or with the same design matrix $\mathbf{X}$, and the number of observations on individuals may become relatively large in some cases. Models that can accommodate the unbalanced nature of longitudinal data and have more parsimonious covariance structures than (1.1) need to be considered. Therefore, to address these problems we consider models for such longitudinal data that contain both individual random effects components and within-individual errors that follow an (autoregressive) AR(1) time series process. The model for individual k is

$$\mathbf{y}_k = \mathbf{X}_k\boldsymbol{\beta} + \mathbf{C}_k\boldsymbol{\tau}_k + \mathbf{u}_k, \qquad k = 1, \ldots, N, \tag{1.2}$$

where $\mathbf{y}_k$ is a $T_k \times 1$ vector of observations, $\mathbf{X}_k$ is the $T_k \times n$ design matrix for the mean vector of individual k, $\boldsymbol{\beta}$ is the $n \times 1$ population fixed effect parameter vector, $\mathbf{C}_k$ is the $T_k \times m$ design matrix for the random effects of individual k, $\boldsymbol{\tau}_k$ is an $m \times 1$ vector of unobservable random effects assumed to be sampled from a multivariate normal distribution with mean $\mathbf{0}$ and $m \times m$ covariance matrix Γ, and $\mathbf{u}_k$ is the $T_k \times 1$ vector of within-individual errors whose components are assumed to follow the AR(1) model

$$u_{k,t} = \phi u_{k,t-1} + \varepsilon_{k,t}, \qquad \varepsilon_{k,t} \sim N(0, \sigma^2). \tag{1.3}$$

It is assumed for individual k that the observations are taken at integer time points $t_{k,1}, \ldots, t_{k,T_k}$, which are not necessarily consecutive, so the "missing data" situation is accommodated. Let $\sigma^2\Omega_k$ denote the covariance matrix of $\mathbf{u}_k$, so for Model (1.2)–(1.3),

$$\Sigma_k = \operatorname{cov}(\mathbf{y}_k) = \mathbf{C}_k\Gamma\mathbf{C}_k' + \sigma^2\Omega_k$$

has a more parsimonious covariance structure than (1.1).

Laird and Ware (1982) considered the random effects model as in (1.2), but their work concentrated on the situation where the within-individual errors are independent, and hence Ω_k was simplified to equal I. This leads to the covariance structure $\Sigma_k = \operatorname{cov}(\mathbf{y}_k) = \mathbf{C}_k\Gamma\mathbf{C}_k' + \sigma^2 I$. A related model in the growth curve setting of the form $\mathbf{y}_k = \mathbf{X}_k\boldsymbol{\beta}_k + \mathbf{u}_k = \mathbf{X}_k\mathbf{B}\mathbf{a}_k + \mathbf{X}_k\boldsymbol{\tau}_k + \mathbf{u}_k$, with $\boldsymbol{\beta}_k = \mathbf{B}\mathbf{a}_k + \boldsymbol{\tau}_k$, was considered by Reinsel (1985). Note that conditional on the random effects $\boldsymbol{\tau}_k$, however, the observations on a given individual are assumed to be independent in the model of Laird and Ware. This suggests that a judicious choice of the random effects terms $\mathbf{C}_k$ in (1.2) may be required to represent adequately the unconditional covariance structure of the $\mathbf{y}_k$ when the errors $\mathbf{u}_k$ are assumed to be independent. Thus it may be appealing to incorporate in addition to the random effects structure in (1.2), the possibility that the individual errors $\mathbf{u}_k$ may be autocorrelated, as in the AR(1) model (1.3).

For the model (1.2) with $\Omega_k = I$, Laird and Ware (1982) discussed in detail how to use the EM algorithm to obtain the maximum likelihood (ML) and restricted maximum likelihood (REML) estimates of the variance–covariance components in Γ and σ^2. More recently, Laird, Lange, and Stram (1987) have continued to examine the application of the EM algorithm to this random effects model. Lindstrom and Bates (1988) examined methods to improve the

* Eric M. Chi is Senior Biostatistician, Merrell Dow Research Institute, Cincinnati, OH 45215. Gregory C. Reinsel is Professor, Department of Statistics, University of Wisconsin, Madison, WI 53706. The authors thank an associate editor for useful comments.

Journal of the American Statistical Association
June 1989, Vol. 84, No. 406, Theory and Methods

computational efficiency of the EM algorithm for these models.

In work related to Model (1.2)–(1.3), LaVange and Helms (1983) used method-of-moments procedures to estimate the time series parameters for the model $\mathbf{y}_k = \mathbf{X}_k\boldsymbol{\beta} + \mathbf{u}_k$, with AR(1) and (moving average) MA(q) models for the process $\mathbf{u}_k$, but without any random effects terms. Jennrich and Schluchter (1986) discussed various types of covariance structures, including random effects models and the AR(1) model separately, and used different algorithms to obtain the ML estimates. Ware (1985) overviewed some special covariance structures, including general multivariate, random effects, and AR(1) models. Azzalini (1986) considered Model (1.2)–(1.3) within the balanced and complete data growth curve setting and obtained closed-form expressions for the ML estimators of $\mathbf{B}$, σ^2, and a transformation of Γ as functions of ϕ. Pantula and Pollock (1985) examined the special case with unbalanced but consecutive data, where $\mathbf{C}_k = \mathbf{1}$ in (1.2), and considered method-of-moments type of estimators for ϕ and the variance components Γ and σ^2. Finally, Jones (1986) discussed the ML estimation of models such as (1.2)–(1.3) through use of a state–space representation and the Kalman filter.

In Section 2 of this article, a score test is first proposed for testing the random effects model (1.2) with $\text{cov}(\mathbf{u}_k) = \sigma^2 I$ against the same model with autocorrelated AR(1) errors for the $\mathbf{u}_k$ as in (1.3). Such a test would be desirable as a check after fitting the "conditional independence" model (1.2) with independent within-individual errors and could serve as a preliminary to fitting Model (1.2)–(1.3) that includes the AR(1) errors. In Section 3 we discuss the ML estimation and, in particular, the scoring approach to obtain the ML estimates for Model (1.2)–(1.3). Estimation of random effects and prediction of future observations for Model (1.2)–(1.3) are discussed in Section 4. A numerical example is presented in Section 5. Some concluding remarks are given in Section 6.

2. THE SCORE TEST FOR AUTOCORRELATION

The score test is originated from Silvey's (1959) Lagrange multiplier approach, which requires only the estimation of the null model, but yields a test asymptotically equivalent to the corresponding likelihood ratio test obtained by overfitting the model. In our context, the appeal of the score test approach for testing $H_0 : \phi = 0$ in (1.3) is that, when the more "standard" random effects model (1.2) is fitted with $\text{cov}(\mathbf{u}_k) = \sigma^2 I$, the score test will provide a relatively simple check for the presence of possible autocorrelation in the errors. We must note that rejecting the null model does not mean the alternate model specified is appropriate, but only suggests that autocorrelation may exist in the $\mathbf{u}_k$'s and the AR(1) model for $\mathbf{u}_k$ represents a reasonable alternative model that is worth consideration.

Let $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_d)' = (\gamma_{11}, \ldots, \gamma_{1m}, \gamma_{22}, \ldots, \gamma_{mm}, \sigma^2)'$ be the vector of all of the variance–covariance components in Model (1.2), where the γ_{ij}'s are the elements above and on the diagonal of the symmetric $m \times m$ matrix Γ, and set $\boldsymbol{\theta}^* = (\boldsymbol{\theta}', \phi)'$. In addition, let $\boldsymbol{\alpha} = (\boldsymbol{\beta}', \boldsymbol{\theta}^{*\prime})' = (\boldsymbol{\beta}', \boldsymbol{\theta}', \phi)'$, and let $\hat{\boldsymbol{\alpha}}_0 = (\hat{\boldsymbol{\beta}}', \hat{\boldsymbol{\theta}}', 0)'$, where $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}$ are the ML estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ in the null model (1.2), which assumes that $\phi = 0$. There are several iterative procedures to obtain the ML estimates $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}$ and the information matrix of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ under H_0, including the scoring procedure in particular. We consider this approach in development of a score statistic for a test of $H_0 : \phi = 0$ in Model (1.2)–(1.3). In Section 3, ML estimation of the alternative model, which includes the autoregressive model specification [AR(1)] for $\mathbf{u}_k$, will be considered.

For the AR(1) model (1.3), the random variables $u_{k,t_{k,j}} - \phi^{\Delta_k(j)} u_{k,t_{k,j-1}}$ are independent with mean 0 and variance $\sigma^2 v_{k,j}$, with $\Delta_k(j) = t_{k,j} - t_{k,j-1}$ for $j = 2, \ldots, T_k$, the times between successive observations for the kth individual, and $v_{k,j} = 1/(1 - \phi^2)$ if $j = 1$ and $v_{k,j} = (1 - \phi^{2\Delta_k(j)})/(1 - \phi^2)$ if $j > 1$. The $\Delta_k(j)$ are positive integers, and for practical applications it is assumed that values of $\Delta_k(j)$ equal to 1, that is, consecutive observations, occur sufficiently often. Hence it follows that Model (1.2)–(1.3) has $\text{cov}(\mathbf{u}_k) = \sigma^2 \Omega_k$, where

$$\Omega_k = (R_k')^{-1} V_k (R_k)^{-1}, \tag{2.1}$$

with $V_k = \text{diag}(v_{k,1}, \ldots, v_{k,T_k})$. R_k' is called the innovator matrix (see Reinsel and Wincek 1987) of Ω_k and has the form

$$R_k' = \begin{bmatrix} 1 & 0 & \cdot & \cdot & 0 \\ -\phi^{\Delta_k(2)} & 1 & & & \cdot \\ 0 & -\phi^{\Delta_k(3)} & 1 & & \cdot \\ \cdot & & \cdot & \cdot & 0 \\ 0 & \cdot & 0 & -\phi^{\Delta_k(T_k)} & 1 \end{bmatrix}.$$

Let l denote the log-likelihood function of $\mathbf{y}_1, \ldots, \mathbf{y}_N$ for Model (1.2)–(1.3), so

$$l = \text{constant} - \frac{1}{2}\sum_{k=1}^{N} \log|\Sigma_k| - \frac{1}{2}\sum_{k=1}^{N} (\mathbf{y}_k - \mathbf{X}_k\boldsymbol{\beta})'\Sigma_k^{-1}(\mathbf{y}_k - \mathbf{X}_k\boldsymbol{\beta}), \tag{2.2}$$

where $\Sigma_k = \text{cov}(\mathbf{y}_k) = \mathbf{C}_k\Gamma\mathbf{C}_k' + \sigma^2\Omega_k$. It follows that when evaluated at $\hat{\boldsymbol{\alpha}}_0$, the score vector $\partial l/\partial\boldsymbol{\alpha}$ has all elements equal to 0 except the derivative with respect to ϕ, denoted by $(\partial l/\partial\phi)_0 = (\partial l/\partial\phi)\,|_{\hat{\boldsymbol{\alpha}}_0}$. The information matrix is $\mathbf{J} = E[-\partial^2 l/\partial\boldsymbol{\alpha}\partial\boldsymbol{\alpha}']$, which has a form block partitioned in accordance with the way $\boldsymbol{\alpha} = (\boldsymbol{\beta}', \boldsymbol{\theta}', \phi)'$ is partitioned as

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{11} & 0 & 0 \\ 0 & \mathbf{J}_{22} & \mathbf{J}_{23} \\ 0' & \mathbf{J}_{23}' & \mathbf{J}_{33} \end{bmatrix}.$$

The score test statistic, denoted by λ, then takes the form

$$\lambda = \left(\frac{\partial l}{\partial\boldsymbol{\alpha}}\bigg|_{\hat{\boldsymbol{\alpha}}_0}\right)'(\mathbf{J}\,|\,_{\hat{\boldsymbol{\alpha}}_0})^{-1}\left(\frac{\partial l}{\partial\boldsymbol{\alpha}}\bigg|_{\hat{\boldsymbol{\alpha}}_0}\right) = \left(\frac{\partial l}{\partial\phi}\right)_0^2 \bigg/ s, \tag{2.3}$$

where $s = \mathbf{J}_{33}^0 - \mathbf{J}_{23}^{0\prime}\mathbf{J}_{22}^{0\,-1}\mathbf{J}_{23}^0$ and the superscript 0 indicates the elements of $\mathbf{J}$ are evaluated at $\hat{\boldsymbol{\alpha}}_0$. Under "regularity" conditions that will ensure asymptotic normality of the score vector, λ is asymptotically chi-squared distributed with 1 df under the null hypothesis. To calculate λ, we first have

$$\frac{\partial l}{\partial \phi} = -\frac{1}{2}\sum_{k=1}^{N} \operatorname{tr}\left(\Sigma_k^{-1}\frac{\partial \Sigma_k}{\partial \phi}\right) + \frac{1}{2}\sum_{k=1}^{N}(\mathbf{y}_k - \mathbf{X}_k\boldsymbol{\beta})'\Sigma_k^{-1}\frac{\partial \Sigma_k}{\partial \phi}\Sigma_k^{-1}(\mathbf{y}_k - \mathbf{X}_k\boldsymbol{\beta}). \tag{2.4}$$

Note that by a matrix inversion formula (Rao 1973, p. 33),

$$\begin{aligned}\Sigma_k^{-1}(\mathbf{y}_k - \mathbf{X}_k\boldsymbol{\beta}) &= \sigma^{-2}(\Omega_k^{-1} - \Omega_k^{-1}\mathbf{C}_k[\mathbf{C}_k'\Omega_k^{-1}\mathbf{C}_k \\ &\quad + \sigma^2\Gamma^{-1}]^{-1}\mathbf{C}_k'\Omega_k^{-1})(\mathbf{y}_k - \mathbf{X}_k\boldsymbol{\beta}) \\ &= \sigma^{-2}\Omega_k^{-1}(\mathbf{y}_k - \mathbf{X}_k\boldsymbol{\beta} - \mathbf{C}_k\bar{\boldsymbol{\tau}}_k),\end{aligned} \tag{2.5}$$

where $\bar{\boldsymbol{\tau}}_k = [\mathbf{C}_k'\Omega_k^{-1}\mathbf{C}_k + \sigma^2\Gamma^{-1}]^{-1}\mathbf{C}_k'\Omega_k^{-1}(\mathbf{y}_k - \mathbf{X}_k\boldsymbol{\beta}) = \Gamma\mathbf{C}_k'\Sigma_k^{-1}(\mathbf{y}_k - \mathbf{X}_k\boldsymbol{\beta}) = E[\boldsymbol{\tau}_k \mid \mathbf{y}_k]$ represents the estimate (posterior mean) of the random effect $\boldsymbol{\tau}_k$, when the model parameter $\boldsymbol{\alpha}$ is known. When evaluated at $\hat{\boldsymbol{\alpha}}_0$, we obtain $\Sigma_k^{-1}(\mathbf{y}_k - \mathbf{X}_k\boldsymbol{\beta})\mid_{\hat{\alpha}_0} = \hat{\sigma}^{-2}(\mathbf{y}_k - \mathbf{X}_k\hat{\boldsymbol{\beta}} - \mathbf{C}_k\hat{\boldsymbol{\tau}}_k) = \hat{\sigma}^{-2}\hat{\mathbf{u}}_k$, where $\hat{\sigma}^2$, $\hat{\boldsymbol{\beta}}$, and $\hat{\boldsymbol{\tau}}_k$ are the estimates of σ^2, $\boldsymbol{\beta}$, and $\boldsymbol{\tau}_k$ in fitting the random effects model with white noise errors and $\hat{\mathbf{u}}_k$ can be viewed as the vector of "residuals" from this model for individual k. In addition, $\partial\Sigma_k/\partial\phi = \sigma^2\partial\Omega_k/\partial\phi$, and since $R_k' = V_k = I$ and $\partial V_k/\partial\phi = 0$ when evaluated at $\hat{\boldsymbol{\alpha}}_0$ (i.e., at $\phi = 0$), it follows from (2.1) that $\partial\Omega_k/\partial\phi \mid_{\hat{\alpha}_0} = L_k' + L_k$, where L_k' is a $T_k \times T_k$ matrix with $(l, l-1)$th element equal to $\delta(\Delta_k(l))$, for $l = 2, \ldots, T_k$, and all other elements equal to 0, where $\delta(\Delta_k(l)) = 1$ if $\Delta_k(l) = 1$ and $\delta(\Delta_k(l)) = 0$ if $\Delta_k(l) > 1$. Using the inversion formula for $\Sigma_k^{-1}\mid_{\hat{\alpha}_0}$ associated with (2.5) we then have $\operatorname{tr}(\Sigma_k^{-1}\partial\Sigma_k/\partial\phi)\mid_{\hat{\alpha}_0} = -2\operatorname{tr}([\mathbf{C}_k'\mathbf{C}_k + \hat{\sigma}^2\hat{\Gamma}^{-1}]^{-1}\mathbf{C}_k'L_k\mathbf{C}_k)$. Thus, from (2.4), we find that

$$\left(\frac{\partial l}{\partial \phi}\right)_0 = \sum_{k=1}^{N}\operatorname{tr}([\mathbf{C}_k'\mathbf{C}_k + \hat{\sigma}^2\hat{\Gamma}^{-1}]^{-1}\mathbf{C}_k'L_k\mathbf{C}_k) + \hat{\sigma}^{-2}\sum_{k=1}^{N}\hat{\mathbf{u}}_k'L_k\hat{\mathbf{u}}_k. \tag{2.6}$$

The second derivatives of l with respect to $\boldsymbol{\theta}^*$ have the form

$$\begin{aligned}\frac{\partial^2 l}{\partial\theta_i^*\partial\theta_j^*} &= \frac{1}{2}\sum_{k=1}^{N}\operatorname{tr}\left(\Sigma_k^{-1}\frac{\partial\Sigma_k}{\partial\theta_j^*}\Sigma_k^{-1}\frac{\partial\Sigma_k}{\partial\theta_i^*}\right) \\ &\quad - \sum_{k=1}^{N}(\mathbf{y}_k - \mathbf{X}_k\boldsymbol{\beta})'\Sigma_k^{-1}\frac{\partial\Sigma_k}{\partial\theta_j^*}\Sigma_k^{-1} \\ &\quad \times \frac{\partial\Sigma_k}{\partial\theta_i^*}\Sigma_k^{-1}(\mathbf{y}_k - \mathbf{X}_k\boldsymbol{\beta}),\end{aligned}$$

with $E[-\partial^2 l/\partial\theta_i^*\partial\theta_j^*] = \frac{1}{2}\sum_{k=1}^{N}\operatorname{tr}[\Sigma_k^{-1}(\partial\Sigma_k/\partial\theta_i^*)\Sigma_k^{-1}(\partial\Sigma_k/\partial\theta_j^*)]$. Thus the elements of $\mathbf{J}_{22}$, $\mathbf{J}_{23}$, and $\mathbf{J}_{33}$ are obtainable from this expression as a function of the Σ_k^{-1} and the derivatives $\partial\Sigma_k/\partial\theta_i^*$. To calculate $\partial\Sigma_k/\partial\theta_i\mid_{\hat{\alpha}_0}$, it can be seen that $\partial\Sigma_k/\partial\gamma_{ij}\mid_{\hat{\alpha}_0} = \mathbf{C}_k(\partial\Gamma/\partial\gamma_{ij})\mathbf{C}_k'$, where $\partial\Gamma/\partial\gamma_{ij}$ has 1 in the (i, j)th and the (j, i)th positions and 0 elsewhere. In addition, it is easily seen that $\partial\Sigma_k/\partial\sigma^2\mid_{\hat{\alpha}_0} = \Omega_k\mid_{\hat{\alpha}_0} = I$. From these results and $\Sigma_k\mid_{\hat{\alpha}_0} = \mathbf{C}_k\hat{\Gamma}\mathbf{C}_k' + \hat{\sigma}^2 I$, $s = \mathbf{J}_{33}^0 - \mathbf{J}_{23}^{0\prime}\mathbf{J}_{22}^{0\,-1}\mathbf{J}_{23}^0$ can easily be calculated.

The score statistic λ can then be calculated, from (2.3) and (2.6), as

$$\lambda = \frac{[\hat{\sigma}^{-2}\sum_{k=1}^{N}\hat{\mathbf{u}}_k'L_k\hat{\mathbf{u}}_k + \sum_{k=1}^{N}\operatorname{tr}([\mathbf{C}_k'\mathbf{C}_k + \hat{\sigma}^2\hat{\Gamma}^{-1}]^{-1}\mathbf{C}_k'L_k\mathbf{C}_k)]^2}{\mathbf{J}_{33}^0 - \mathbf{J}_{23}^{0\prime}\mathbf{J}_{22}^{0\,-1}\mathbf{J}_{23}^0}. \tag{2.7}$$

It should be noted that $\hat{\mathbf{u}}_k'L_k\hat{\mathbf{u}}_k = \Sigma_t'\,\hat{u}_{k,t}\hat{u}_{k,t-1}$, where Σ_t' denotes summation over those observation times $t_{k,j}$ for which an observation at the previous consecutive time $t_{k,j-1} = t_{k,j} - 1$ is also available. Hence $(1/T^*)\sum_{k=1}^{N}\hat{\mathbf{u}}_k'L_k\hat{\mathbf{u}}_k = (1/T^*)\sum_{k=1}^{N}\Sigma_t'\,\hat{u}_{k,t}\hat{u}_{k,t-1} = c^*(1)$, with $T^* = \sum_{k=1}^{N}\sum_{l=2}^{T_k}\delta(\Delta_k(l))$, has the form of a lag one sample autocovariance for the $\hat{u}_{k,t}$ pooled over the N individuals. The numerator in (2.7) thus has the form of the square of a pooled sample autocorrelation at lag one of the $\hat{u}_{k,t}$, with a "correction" for bias due to the use of the estimates $\hat{\boldsymbol{\tau}}_k$ in the formation of the $\hat{\mathbf{u}}_k$. It is interesting to note that, from the score test approach, in the presence of the random effects terms in model (1.2) the appropriate residuals $\hat{\mathbf{u}}_k$ upon which the test for autocorrelation is based are $\hat{\mathbf{u}}_k = \mathbf{y}_k - \mathbf{X}_k\hat{\boldsymbol{\beta}} - \mathbf{C}_k\hat{\boldsymbol{\tau}}_k$, where $\hat{\boldsymbol{\tau}}_k$ is the empirical Bayes estimator of the random effects $\boldsymbol{\tau}_k$ (under the null model). A large value of λ, compared with a χ^2 distribution with 1 df, would suggest autocorrelation in the $\mathbf{u}_k$'s and might lead to consideration of the ML estimation of the alternate model that includes the AR(1) errors model (1.3). Although the score test statistic (2.7) is valid for irregular patterns of observation times, the power of the test will tend to be best when observations are consecutive.

We remark that in the particular case of a simple random effects model with $\mathbf{C}_k = \mathbf{1}$, a $T_k \times 1$ column of ones, the numerator of the score statistic (2.7) reduces to the square of $T^*r^*(1) + \hat{\Gamma}\sum_{k=1}^{N}[T_k^*/(\hat{\sigma}^2 + T_k\hat{\Gamma})]$, where $T_k^* = \sum_{l=2}^{T_k}\delta(\Delta_k(l))$, and $r^*(1) = c^*(1)/\hat{\sigma}^2$ has the form of a pooled lag one sample autocorrelation of the $\hat{u}_{k,t}$. In addition, the $\hat{\mathbf{u}}_k$ are given by $\hat{\mathbf{u}}_k = \mathbf{y}_k - \mathbf{X}_k\hat{\boldsymbol{\beta}} - \mathbf{1}\hat{\tau}_k$, where $\hat{\tau}_k = [T_k\hat{\Gamma}/(T_k\hat{\Gamma} + \hat{\sigma}^2)]\tau_k^*$ and $\tau_k^* = T_k^{-1}\mathbf{1}'(\mathbf{y}_k - \mathbf{X}_k\hat{\boldsymbol{\beta}})$ is the average of the deviations from the estimated mean response for the kth individual. Furthermore, in the special case of balanced consecutive data in the simple random effects model, with all $T_k = T$, the score statistic reduces to $\lambda = NT(T-1)[r^*(1) + \hat{P}]^2/\{(T-2)[1 - 2\hat{P}/(T-1)]\}$, where $\hat{P} = \hat{\Gamma}/(T\hat{\Gamma} + \hat{\sigma}^2)$.

3. SCORING METHOD FOR MAXIMUM LIKELIHOOD ESTIMATION IN THE RANDOM EFFECTS MODEL WITH AR(1) ERRORS

If the score test developed in Section 2 rejects the null model, then the alternate model [(1.2)–(1.3)] may be fitted by maximum likelihood. The model implies that $\operatorname{cov}(\mathbf{u}_k) = \sigma^2\Omega_k$, where $\Omega_k = (R_k')^{-1}V_k(R_k)^{-1}$ as given in (2.1), and thus $\Sigma_k = \operatorname{cov}(\mathbf{y}_k) = \mathbf{C}_k\Gamma\mathbf{C}_k' + \sigma^2\Omega_k$. The EM algorithm approach to ML estimation, such as that of Laird and Ware (1982), is not so appealing in the presence of a time series model such as (1.3) for the errors $u_{k,t}$, since the M step would not yield an explicit closed form of the ML estimator in a one-step linear calculation. Thus the scoring method for estimation of this model will now be considered.

To apply the scoring procedure, we consider the log-likelihood of $\mathbf{y}_1, \ldots, \mathbf{y}_N$ as given by (2.2). If ML estimates $\hat{\boldsymbol{\theta}}^* = (\hat{\boldsymbol{\theta}}', \hat{\phi})'$ of $\boldsymbol{\theta}^* = (\boldsymbol{\theta}', \phi)'$ were available, the ML estimate $\hat{\boldsymbol{\beta}}$ can be obtained as $\hat{\boldsymbol{\beta}} = (\sum_{k=1}^N \mathbf{X}_k'\hat{\Sigma}_k^{-1}\mathbf{X}_k)^{-1} \sum_{k=1}^N \mathbf{X}_k'\hat{\Sigma}_k^{-1}\mathbf{y}_k$, where $\hat{\Sigma}_k = \mathbf{C}_k\hat{\Gamma}\mathbf{C}_k' + \hat{\sigma}^2\hat{\Omega}_k$. Thus in the iteration procedure, we will first obtain $\hat{\boldsymbol{\theta}}^*$, then $\hat{\boldsymbol{\beta}}$, and these are used as the estimates of $\boldsymbol{\theta}^*$ and $\boldsymbol{\beta}$ to start the next iteration. It will also be required that the estimate $\hat{\Gamma}$ of the random effects covariance matrix be nonnegative definite. Reparameterization of $\Gamma = F'F$ by the Cholesky decomposition (Lindstrom and Bates 1988), where F is an upper triangular matrix, will result in a nonnegative-definite matrix Γ. Thus, instead of Γ, we treat F as the parameters.

Letting $\boldsymbol{\eta} = (f_{11}, f_{12}, \ldots, f_{1m}, f_{22}, \ldots, f_{2m}, f_{33}, \ldots, f_{mm}, \sigma^2, \phi)'$, we need to calculate the score vector $\mathbf{s} = \partial l/\partial\boldsymbol{\eta}$ and the expectation of the negative Hessian matrix $\mathbf{J}^* = E[-\partial^2 l/\partial\boldsymbol{\eta}\partial\boldsymbol{\eta}']$. With the current estimates $\hat{\boldsymbol{\eta}}^{(h)}$ and $\hat{\boldsymbol{\beta}}^{(h)}$ at the hth iteration step, one iteration of the scoring procedure for estimation of $\boldsymbol{\eta}$ is then

$$\hat{\boldsymbol{\eta}}^{(h+1)} = \hat{\boldsymbol{\eta}}^{(h)} + \hat{\mathbf{J}}_*^{(h)-1}\hat{\mathbf{s}}^{(h)}, \tag{3.1}$$

where $\hat{\mathbf{s}}^{(h)}$ and $\hat{\mathbf{J}}_*^{(h)}$ denote $\mathbf{s}$ and $\mathbf{J}_*$ evaluated at $\hat{\boldsymbol{\eta}}^{(h)}$ and $\hat{\boldsymbol{\beta}}^{(h)}$. Then $\hat{\boldsymbol{\beta}}^{(h+1)}$ is obtained as

$$\hat{\boldsymbol{\beta}}^{(h+1)} = \left(\sum_{k=1}^N \mathbf{X}_k'\hat{\Sigma}_k^{(h+1)-1}\mathbf{X}_k\right)^{-1} \sum_{k=1}^N \mathbf{X}_k'\hat{\Sigma}_k^{(h+1)-1}\mathbf{y}_k, \tag{3.2}$$

where $\hat{\Sigma}_k^{(h+1)} = \mathbf{C}_k\hat{\Gamma}^{(h+1)}\mathbf{C}_k' + \hat{\sigma}^{2(h+1)}\hat{\Omega}_k^{(h+1)}$ with $\hat{\Gamma}^{(h+1)} = \hat{F}^{(h+1)'}\hat{F}^{(h+1)}$. The ML estimates $\hat{\boldsymbol{\beta}}_M$ and $\hat{\boldsymbol{\eta}}_M$ are obtained by iteration of (3.1) and (3.2), and upon convergence we use $\text{cov}(\hat{\boldsymbol{\beta}}_M) = (\sum_{k=1}^N \mathbf{X}_k'\hat{\Sigma}_k^{-1}\mathbf{X}_k)^{-1}$ and $\text{cov}(\hat{\boldsymbol{\eta}}_M) = \hat{\mathbf{J}}_*^{-1}$ as the covariance matrices of the ML estimates.

To start the iterations, $\hat{\boldsymbol{\eta}}^{(0)} = (\hat{f}_{11}, \ldots, \hat{f}_{mm}, \hat{\sigma}^2, 0)'$ are used as the initial estimates, where the $\hat{f}_{ij}$'s are obtained from the $\hat{\gamma}_{ij}$'s by Cholesky decomposition and the $\hat{\gamma}_{ij}$'s and $\hat{\sigma}^2$ are the ML estimates of the γ_{ij} and σ^2 in the random effects white noise errors model.

To calculate $\mathbf{s}$ and $\mathbf{J}_*$, we have for any η_i, η_j,

$$\frac{\partial l}{\partial \eta_i} = -\frac{1}{2}\sum_{k=1}^N \text{tr}\left(\Sigma_k^{-1}\frac{\partial \Sigma_k}{\partial \eta_i}\right) + \frac{1}{2}\sum_{k=1}^N (\mathbf{y}_k - \mathbf{X}_k\boldsymbol{\beta})'\Sigma_k^{-1}\frac{\partial \Sigma_k}{\partial \eta_i}\Sigma_k^{-1}(\mathbf{y}_k - \mathbf{X}_k\boldsymbol{\beta})$$

and

$$E\left[-\frac{\partial^2 l}{\partial \eta_i \partial \eta_j}\right] = \frac{1}{2}\sum_{k=1}^N \text{tr}\left(\Sigma_k^{-1}\frac{\partial \Sigma_k}{\partial \eta_i}\Sigma_k^{-1}\frac{\partial \Sigma_k}{\partial \eta_j}\right).$$

In particular, since $\Sigma_k = \mathbf{C}_k F'F\mathbf{C}_k' + \sigma^2\Omega_k$,

$$\frac{\partial \Sigma_k}{\partial f_{ij}} = \mathbf{C}_k\frac{\partial F'}{\partial f_{ij}}F\mathbf{C}_k' + \mathbf{C}_k F'\frac{\partial F}{\partial f_{ij}}\mathbf{C}_k', \qquad i \le j,$$

$$\frac{\partial \Sigma_k}{\partial \sigma^2} = \Omega_k,$$

and

$$\frac{\partial \Sigma_k}{\partial \phi} = \sigma^2\frac{\partial \Omega_k}{\partial \phi} = \sigma^2\left[-R_k'^{-1}\frac{\partial R_k'}{\partial \phi}R_k'^{-1}V_kR_k^{-1} + R_k'^{-1}\frac{\partial V_k}{\partial \phi}R_k^{-1} - R_k'^{-1}V_kR_k^{-1}\frac{\partial R_k}{\partial \phi}R_k^{-1}\right].$$

For computation, to obtain the inverse of the $T_k \times T_k$ matrix Σ_k, a matrix inversion formula (Rao 1973, p. 33),

$$\Sigma_k^{-1} = \sigma^{-2}\Omega_k^{-1} - \sigma^{-2}\Omega_k^{-1}\mathbf{C}_k \times [\sigma^2 I + F'F(\mathbf{C}_k'\Omega_k^{-1}\mathbf{C}_k)]^{-1}F'F\mathbf{C}_k'\Omega_k^{-1}, \tag{3.3}$$

is used, which involves only inversion of Ω_k and an $m \times m$ matrix. From (2.1), the relation $\Omega_k^{-1} = R_kV_k^{-1}R_k'$ is also used in the computations. Using these results, the calculations in (3.1) to obtain $\hat{\boldsymbol{\eta}}$ can easily be performed.

In addition, these results can be used to express $\hat{\boldsymbol{\beta}}$ more conveniently in terms of variables in the transformed model corresponding to (1.2) with white noise errors, $\mathbf{y}_k^* = \mathbf{X}_k^*\boldsymbol{\beta} + \mathbf{C}_k^*\boldsymbol{\tau}_k + \boldsymbol{\epsilon}_k$, where $\mathbf{y}_k^* = V_k^{-1/2}R_k'\mathbf{y}_k$, $\mathbf{X}_k^* = V_k^{-1/2}R_k'\mathbf{X}_k$, and $\mathbf{C}_k^* = V_k^{-1/2}R_k'\mathbf{C}_k$. Thus it follows that $\hat{\boldsymbol{\beta}}$ can be written as $\hat{\boldsymbol{\beta}} = (\sum_{k=1}^N \mathbf{X}_k^{*\prime}\Sigma_k^{*-1}\mathbf{X}_k^*)^{-1}\sum_{k=1}^N \mathbf{X}_k^{*\prime}\Sigma_k^{*-1}\mathbf{y}_k^*$, where $\Sigma_k^{*-1} = (\mathbf{C}_k^*\Gamma\mathbf{C}_k^{*\prime} + \sigma^2 I)^{-1} = \sigma^{-2}(I - \mathbf{C}_k^*[\sigma^2 I + \Gamma(\mathbf{C}_k^{*\prime}\mathbf{C}_k^*)]^{-1}\Gamma\mathbf{C}_k^{*\prime})$ is the inverse of the covariance matrix of $\mathbf{y}_k^*$ in the (transformed) random effects model with white noise errors.

The value of -2 times the log-likelihood, denoted by $-2l^{(h)}$, is calculated at $\hat{\boldsymbol{\eta}}^{(h)}$ for each iteration. It is required that $-2l^{(h+1)} \le -2l^{(h)}$ for each iteration, but sometimes the increment $\hat{\mathbf{J}}_*^{(h)-1}\hat{\mathbf{s}}^{(h)}$ is too large. Then (3.1) is modified as $\hat{\boldsymbol{\eta}}^{(h+1)} = \hat{\boldsymbol{\eta}}^{(h)} + \rho^{(h)}\hat{\mathbf{J}}_*^{(h)-1}\hat{\mathbf{s}}^{(h)}$, where $\rho^{(h)} \in (0, 1)$ is found such that $-2l^{(h+1)} \le -2l^{(h)}$.

One useful feature of the scoring method in (3.1) is that $\hat{\mathbf{J}}_*^{(h)}$ is always guaranteed to be nonnegative definite. From the experiences of data fitting, however, sometimes $\hat{\mathbf{J}}_*^{(h)}$ tends to have one or more eigenvalues very close to 0, which probably means there is an overparameterization problem. This may occur when the matrix Γ is singular, indicating a rank deficiency in the random effects structure. In such a case, after suitable arrangement of the random effects, Γ has the factorization $\Gamma = F_*'F_*$, where F_* is upper triangular but now is of dimension $r \times m$, $r < m$; that is, $f_{ij} = 0$ for $i > r$. Thus, when appropriate, the restrictions $f_{ij} = 0$ for $i > r$ in the random effects model with AR(1) errors can easily be incorporated in the estimation procedure (3.1).

The scoring method (3.1) to obtain the ML estimates of $\boldsymbol{\eta}$ can readily be modified to obtain the REML's of these parameters. Following Harville (1974), the restricted log-likelihood function to be maximized is the function l plus the additional factor $-(\frac{1}{2})\log|\sum_{k=1}^N \mathbf{X}_k'\Sigma_k^{-1}\mathbf{X}_k|$. Then, for example, the modified score vector is $\mathbf{s} = \partial l/\partial\boldsymbol{\eta}$ plus the additional term whose ith component is $-(\frac{1}{2})\partial(\log|\sum_{k=1}^N \mathbf{X}_k'\Sigma_k^{-1}\mathbf{X}_k|)/\partial\eta_i = \frac{1}{2}\text{tr}[(\sum_{k=1}^N \mathbf{X}_k'\Sigma_k^{-1}\mathbf{X}_k)^{-1}\sum_{k=1}^N \mathbf{X}_k'\Sigma_k^{-1}(\partial\Sigma_k/\partial\eta_i)\Sigma_k^{-1}\mathbf{X}_k]$. Similar modifications are involved in the Hessian matrix, which lead to modified scoring equations of the same form as (3.1) for the REML of η.

4. EMPIRICAL BAYES ESTIMATES OF RANDOM EFFECTS AND PREDICTION OF FUTURE VALUES

The estimation of the individual random effects τ_k in Model (1.2)–(1.3) may sometimes also be of interest, especially for problems of prediction and selection. The minimum mean squared error (MSE) estimator of τ_k is the conditional mean of τ_k, given y_k, which is $\tilde{\tau}_k = \Gamma C_k' \Sigma_k^{-1}(y_k - X_k\beta)$. Using (3.3), $\tilde{\tau}_k$ can be expressed as

$$\tilde{\tau}_k = \tau_k^* - W_k(W_k + \Gamma)^{-1}\tau_k^*, \quad (4.1)$$

where $W_k = \sigma^2(C_k'\Omega_k^{-1}C_k)^{-1} = \sigma^2(C_k^{*\prime}C_k^*)^{-1}$, and $\tau_k^* = (C_k'\Omega_k^{-1}C_k)^{-1}C_k'\Omega_k^{-1}(y_k - X_k\beta) = (C_k^{*\prime}C_k^*)^{-1}C_k^{*\prime}(y_k^* - X_k^*\beta)$ is the individual generalized least squares estimator of τ_k under the AR(1) errors model when β is assumed known and τ_k is treated as fixed. The MSE matrix of $\tilde{\tau}_k$ is given by

$$E[(\tilde{\tau}_k - \tau_k)(\tilde{\tau}_k - \tau_k)'] = W_k - W_k(W_k + \Gamma)^{-1}W_k. \quad (4.2)$$

In practice, the parameters β and η of Model (1.2)–(1.3) are unknown, so ML estimates $\hat{\beta}_M$ and $\hat{\eta}_M$ are used in (4.1) to obtain the empirical Bayes estimator

$$\hat{\tau}_k = \tau_k^+ - \hat{W}_k(\hat{W}_k + \hat{\Gamma})^{-1}\tau_k^+, \quad (4.3)$$

where $\hat{W}_k = \hat{\sigma}^2(C_k'\hat{\Omega}_k^{-1}C_k)^{-1}$ and τ_k^+ is τ_k^* evaluated at $\hat{\beta}_M$ and $\hat{\eta}_M$. The MSE matrix of $\hat{\tau}_k$ may be approximated by the expression in (4.2), but this expression does not account for errors due to estimation of the unknown parameters β and η and hence it will tend to underestimate the actual MSE matrix for smaller sample sizes N. The actual MSE matrix and more accurate approximations to the MSE matrix of $\hat{\tau}_k$ that account for parameter estimation errors have been considered for special cases of the general model [(1.2)–(1.3)] by Reinsel (1985). Similar approaches may be employed for the general case to obtain a better approximation to the MSE matrix than (4.2) for smaller sample sizes N, but this will not be considered in this article.

Prediction of future values of an individual within the longitudinal data context may also be of interest and has been considered by several authors, including Lee and Geisser (1975), Geisser (1981), and Reinsel (1984). Thus we consider the situation where observations y_{01} are available for a further individual over the first portion of time, and we are interested in the prediction of the values y_{02} for this individual over the next (future) portion based on Model (1.2)–(1.3). The model for $y_0 = (y_{01}', y_{02}')'$ can be written as $y_0 = X_0\beta + C_0\tau_0 + u_0$, where

$$X_0 = \begin{bmatrix} X_{01} \\ X_{02} \end{bmatrix}, \quad C_0 = \begin{bmatrix} C_{01} \\ C_{02} \end{bmatrix}, \quad u_0 = \begin{bmatrix} u_{01} \\ u_{02} \end{bmatrix},$$

with $\text{cov}(u_0) = \sigma^2\Omega_0$, and Ω_0 is partitioned into blocks Ω_{ij} ($i, j = 1, 2$). Then $\text{cov}(y_0) = \Sigma_0 = C_0\Gamma C_0' + \sigma^2\Omega_0$, which is also partitioned into blocks $\Sigma_{ij} = C_{0i}\Gamma C_{0j}' + \sigma^2\Omega_{ij}$. The minimum MSE predictor $\hat{y}_{02}$ of y_{02}, based on y_{01}, is given by

$$\hat{y}_{02} = E[y_{02} \mid y_{01}] = X_{02}\beta + C_{02}\hat{\tau}_0 + \hat{u}_{02}, \quad (4.4)$$

where $\hat{\tau}_0 = E[\tau_0 \mid y_{01}]$ and $\hat{u}_{02} = E[u_{02} \mid y_{01}]$. Thus $\hat{\tau}_0 = \tau_0^* - W_{01}(W_{01} + \Gamma)^{-1}\tau_0^*$, with $W_{01} = \sigma^2(C_{01}'\Omega_{11}^{-1}C_{01})^{-1}$ and $\tau_0^* = (C_{01}'\Omega_{11}^{-1}C_{01})^{-1}C_{01}'\Omega_{11}^{-1}(y_{01} - X_{01}\beta)$ has the same form as in (4.1). In addition, it can be shown that

$$\hat{u}_{02} = \Omega_{21}\Omega_{11}^{-1}(y_{01} - X_{01}\beta - C_{01}\hat{\tau}_0), \quad (4.5)$$

where $\Omega_{21}\Omega_{11}^{-1}$ has all columns equal to $\mathbf{0}$ except the last column, which has elements of the form $\phi^{(t_{0,j} - t_{0,T_{01}})}$ ($j = T_{01} + 1, \ldots, T_0$), where T_{01} denotes the number of observations in y_{01}. That is, the elements of $\hat{u}_{02}$ are simply $\hat{u}_{0,j} = \phi^{(t_{0,j} - t_{0,T_{01}})}\hat{u}_{0,T_{01}}$, where $\hat{u}_{0,T_{01}}$ is the last element of the vector of residuals $\hat{u}_{01} = y_{01} - X_{01}\beta - C_{01}\hat{\tau}_0$, and hence the elements of $\hat{u}_{02}$ represent the usual predictions from an AR(1) model based on "residual" values $\hat{u}_{01}$. The MSE matrix of the predictor (4.4) is given by

$$\begin{aligned} E[(y_{02} - \hat{y}_{02})(y_{02} - \hat{y}_{02})'] &= \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \\ &= \sigma^2(\Omega_{22} - \Omega_{21}\Omega_{11}^{-1}\Omega_{12}) + (C_{02} - \Omega_{21}\Omega_{11}^{-1}C_{01}) \\ &\quad \times (W_{01} - W_{01}(W_{01} + \Gamma)^{-1}W_{01}) \\ &\quad \times (C_{02} - \Omega_{21}\Omega_{11}^{-1}C_{01})'. \end{aligned} \quad (4.6)$$

In practice, ML estimates of unknown parameters β and η are used in the expressions of (4.4) and (4.5), and an approximation for the MSE matrix of the resulting predictor is given by (4.6); although, again, this expression does not take into account the effects of the use of estimates of the true parameter values on the prediction MSE matrix.

5. A NUMERICAL EXAMPLE

We now discuss a numerical example that involves data from Zerbe (1979). A random effects model analysis of these data was considered by Reinsel (1984). In the study, standard glucose tolerance tests were administered to 20 obese patients, and plasma inorganic phosphate (mg/dl) measurements determined from blood samples were obtained for each patient at 0, .5, 1, 1.5, 2, 3, 4, and 5 hours after the glucose challenge. From the plot of the averages of plasma inorganic phosphate measurements over these 20 patients at each time point (Fig. 1), it is reasonable to model the mean response by a piecewise linear function with change point at 2 hours. Thus the design matrix for the mean response of all 20 patients has the form

$$X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & .5 & 1 & 1.5 & 2 & 2 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 2 & 3 \end{bmatrix}'.$$

To model the between-individual variation (random effects), we start by assuming that only the intercept is random, and thus $C_k = \mathbf{1}$ for all k. This model with white noise errors,

$$y_k = X\beta + \mathbf{1}\tau_k + u_k, \quad \tau_k \sim N(0, \Gamma), \quad u_k \sim N(0, \sigma^2 I), \quad k = 1, \ldots, 20, \quad (5.1)$$

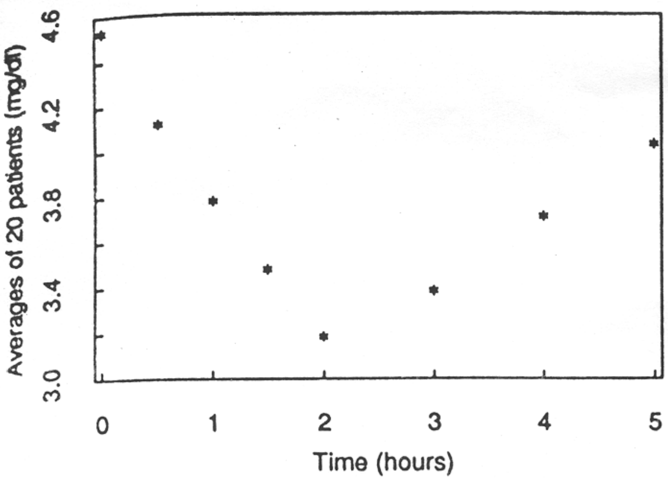

Figure 1. The Averages of Plasma Inorganic Phosphate Measurements Over 20 Patients

was fitted by the scoring method. The resulting ML estimates of σ^2 and F, where F is such that $\Gamma = F^2$, are $\hat{\sigma}^2 = .1515$ and $\hat{F} = .6013$, and the value of -2 times the log-likelihood corresponding to this fitted model is $-2l = 212.154$.

The score test statistic for autocorrelation of the $\mathbf{u}_k$ in (5.1) gave a value of $\lambda = 45.85$, which is highly significant compared with a χ_1^2 distribution. This score test statistic value was obtained by treating the eight time points as each being one "time" unit apart, whereas if one half an hour is taken as the basic time unit and the actual times of the observations are used, the score test statistic value was 26.73, which is still significant but less so than the former value. In addition, when the random effects model with AR(1) errors was fitted, with all points treated as each being one unit apart, a better fit is obtained than with one half an hour treated as the basic time unit. Hence the eight time points will be taken as each being one unit apart in our further analysis.

The score test suggests the possible existence of autocorrelation in the within-individual errors $\mathbf{u}_k$'s. The alternate model

$$\mathbf{y}_k = \mathbf{X}\boldsymbol{\beta} + \mathbf{1}\tau_k + \mathbf{u}_k, \qquad \tau_k \sim N(0, \Gamma),$$
$$k = 1, \ldots, 20, \quad (5.2)$$

with $\mathbf{u}_k$ an AR(1) process, was fitted. The ML estimates, along with the value of -2 times the log-likelihood for this model, are given in Table 1. Another model, with more general random effects and white noise errors,

$$\mathbf{y}_k = \mathbf{X}\boldsymbol{\beta} + \mathbf{X}\boldsymbol{\tau}_k + \mathbf{u}_k, \qquad \boldsymbol{\tau}_k \sim N(\mathbf{0}, \Gamma),$$
$$\mathbf{u}_k \sim N(\mathbf{0}, \sigma^2 I), \quad k = 1, \ldots, 20, \quad (5.3)$$

was also fitted, and results are given in Table 2.

The score test statistic for autocorrelation of Model (5.3) was found to be 2.50, which means this model does explain most of the autocorrelation in Model (5.1), but when compared with the simple random effects ($\mathbf{C}_k = \mathbf{1}$) with AR(1) errors model (5.2), the latter has fewer parameters and smaller value of -2 times the log-likelihood. Hence it is preferred over Model (5.3). It is also noted that the likelihood ratio statistic to test for autocorrelation of the $\mathbf{u}_k$ in Model (5.1), as given by the difference in values of -2 times the log-likelihood between Models (5.2) and (5.1), is $212.15 - 161.12 = 51.03$, which is relatively close to the score test statistic value of 45.85.

Table 1. ML Estimation Results for the Random Effects ($C_k = 1$) Model With AR(1) Errors, (5.2), Where F Is Such That $\Gamma = F^2$

$\hat{\beta}$	$\hat{F}$	$\hat{\sigma}^2$	$\hat{\phi}$	$-2l$	Number of parameters
4.506					
(.158)					
−.677	.5142	.1265	.6973	161.116	6
(.068)	(.1291)	(.0157)	(.1039)		
.286					
(.042)					

The model that combines the general random effects structure of (5.3), with $\mathbf{C}_k = \mathbf{X}$, and the AR(1) errors structure of (5.2) was also considered. Thus the model

$$\mathbf{y}_k = \mathbf{X}\boldsymbol{\beta} + \mathbf{X}\boldsymbol{\tau}_k + \mathbf{u}_k, \qquad \boldsymbol{\tau}_k \sim N(\mathbf{0}, \Gamma),$$
$$k = 1, \ldots, N, \quad (5.4)$$

with the $\mathbf{u}_k$ an AR(1) process, was also estimated by the method of Section 3. The value of -2 times the log-likelihood was found to be 159.09, which does not improve much over the model (5.2) with the value of -2 times the log-likelihood being 161.12. The estimate of ϕ was .3957 with standard deviation .2211. These results indicate that the "extra" random effects terms are not needed. This is supported by the fact that, excluding $\hat{f}_{11}$, most of the elements of $\hat{F}$ for this model were relatively small compared with their respective standard errors.

Neither Model (5.3) nor Model (5.4) offers any substantial improvement over Model (5.2); thus Model (5.2) is our preferred choice. The fitted values from the random effects AR(1) errors model (5.2) for all 20 patients, $\hat{\mathbf{y}}_k = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{1}\hat{\tau}_k$ $(k = 1, \ldots, 20)$, where $\hat{\tau}_k$ is the empirical Bayes estimator (4.3) with all of the parameters replaced by their ML estimates, are plotted in Figure 2. The "predicted" values from this model for all 20 patients, $\hat{\mathbf{y}}_k = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{1}\hat{\tau}_k + \hat{\mathbf{u}}_k(1)$ $(k = 1, \ldots, 20)$, where $\hat{\mathbf{u}}_k(1)$ denotes the vector of one-step-ahead predictions for $\mathbf{u}_k$ based on the previous $\hat{u}_{k,t}$'s and using the AR(1) model, are also plotted in Figure 2. Figure 2 clearly indicates the improvement in predictions from the model by inclusion of the prediction components associated with the AR(1) noise model.

Table 2. ML Estimation Results for the "General" Random Effects ($C_k = X$) Model With White Noise Errors, (5.3), Where F Is the Cholesky Decomposition of Γ Such That $\Gamma = F'F$

$\hat{\beta}$	$\hat{F}$			$\hat{\sigma}^2$	$-2l$	Number of parameters
4.491	.763	−.156	−.032			
(.176)	(.129)	(.072)	(.048)			
−.678		.248	−.070	.0667	162.177	10
(.074)		(.049)	(.050)	(.0094)		
.292			.163			
(.046)			(.033)			

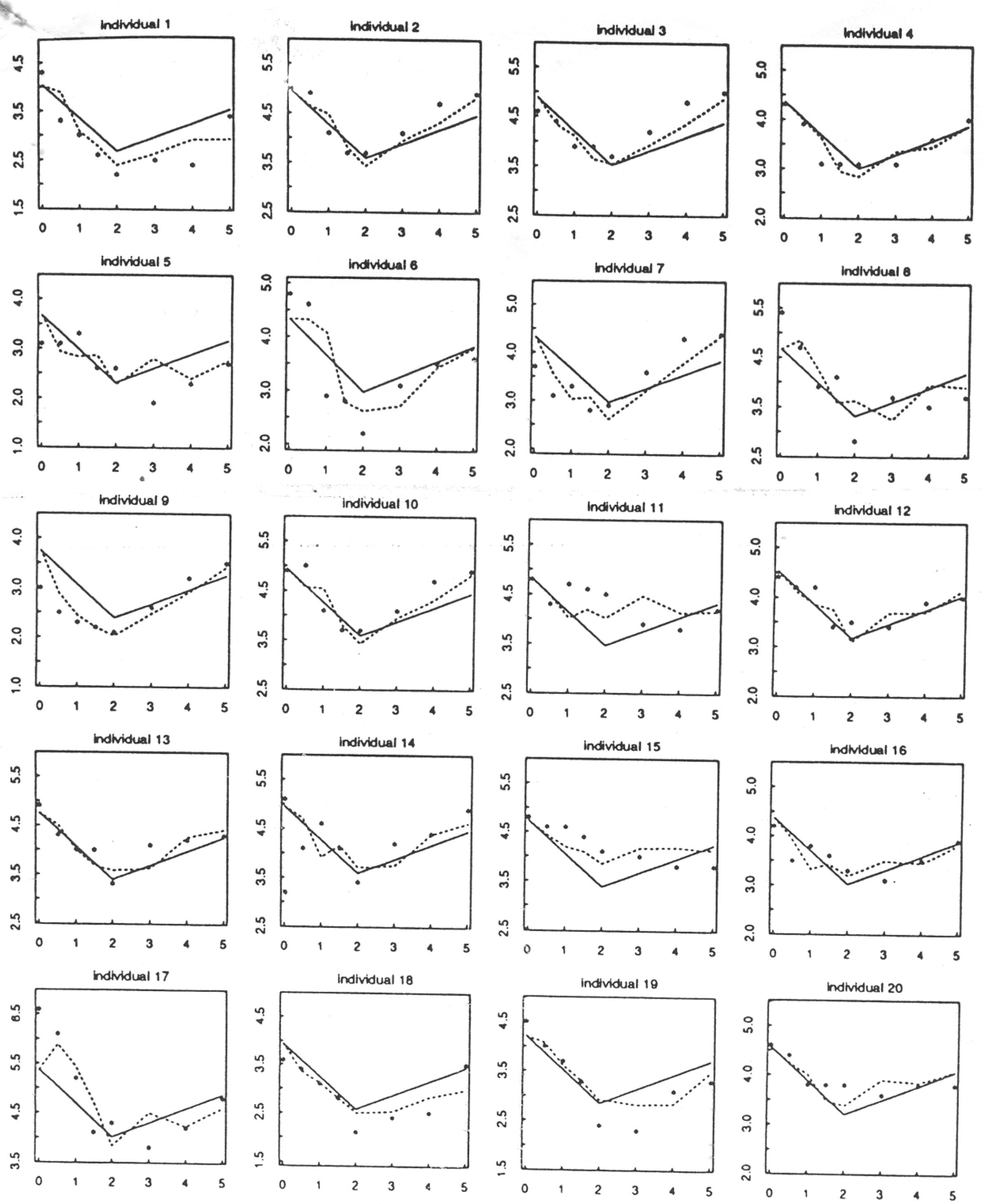

Figure 2. Plasma Inorganic Phosphate Measurement Data With Fitted and Predicted Values Against Time (hours) for 20 Patients: •, Original Data; ——, Fitted Values; ---, Predicted Values.

6. CONCLUDING REMARKS

For situations in which the numbers of observations T_i on the individuals are relatively similar and not too large, inclusion of a sufficient number of random effects terms in the model with white noise errors may be able to represent the serial correlations among the measurements on each individual. However, modeling of autocorrelations through use of time series models for the errors, such as an AR(1) model, in addition to the inclusion of random effects terms, may be more appropriate and may lead to a more proper representation of the correlation structure. Use of an AR(1) errors model can also have the effect of "reducing" the number of random effects needed in the model. Thus methods for determination of the best combination of time series model for errors and random effects terms for the model is an important topic that deserves further consideration. Generalizations of the AR(1) model to other time series models, such as higher order AR or MA models, may also be worth consideration, and the methods of this article can be directly modified to include these models based on the approach of Wincek and Reinsel (1986). Selection of an appropriate form of time series model for the errors may be based on a sequence of likelihood ratio tests, on use of information criteria such as the Akaike information criterion, or on cross-validation prediction methods to compare time series models more complicated than AR(1). In conclusion, in the modeling of longitudinal data we favor consideration and inclusion of the AR(1) or other time series specification, in addition to possible random effects, because it may lead to a more parsimonious correlation model and has the potential to provide an accurate representation for the correlation structure among repeated measurements.

[*Received March 1988. Revised October 1988.*]

REFERENCES

Azzalini, A. (1987), "Growth Curves Analysis for Patterned Covariance Matrices," in *New Perspectives in Theoretical and Applied Statistics*, eds. M. L. Puri, J. P. Vilaplana, and W. Wertz, New York: John Wiley, pp. 63–73.

Geisser, S. (1981), "Sample Reuse Procedures for Prediction of the Unobserved Portion of a Partially Observed Vector," *Biometrika*, 68, 243–250.

Grizzle, J. E., and Allen, D. M. (1969), "Analysis of Growth and Dose Response Curves," *Biometrics*, 25, 357–381.

Harville, D. A. (1974), "Bayesian Inference for Variance Components Using Only Error Contrasts," *Biometrika*, 61, 383–385.

Jennrich, R. I., and Schluchter, M. D. (1986), "Unbalanced Repeated Measures Models With Structural Covariance Matrices," *Biometrics*, 42, 805–820.

Jones, R. H. (1986), "Random Effects and the Kalman Filter," in *Proceedings of the Business and Economic Statistics Section, American Statistical Association*, pp. 69–75.

Laird, N. M., Lange, N., and Stram, D. O. (1987), "Maximum Likelihood Computations With Repeated Measures: Application of the EM Algorithm," *Journal of the American Statistical Association*, 82, 97–105.

Laird, N. M., and Ware, J. H. (1982), "Random-Effects Models for Longitudinal Data," *Biometrics*, 38, 963–974.

LaVange, L. M., and Helms, R. W. (1983), "The Analysis of Incomplete Longitudinal Data With Time Series Covariance Structures," unpublished paper presented at the Joint Statistical Meetings, Toronto.

Lee, J. C., and Geisser, S. (1975), "Applications of Growth Curve Prediction," *Sankhyā*, Ser. A, 37, 239–256.

Lindstrom, M. J., and Bates, D. M. (1988), "Newton–Raphson and EM Algorithms for Linear Mixed-Effects Models for Repeated-Measures Data," *Journal of the American Statistical Association*, 83, 1014–1022.

Pantula, S. G., and Pollock, K. H. (1985), "Nested Analysis of Variance With Autocorrelated Errors," *Biometrics*, 41, 909–920.

Potthoff, R. F., and Roy, S. N. (1964), "A Generalized Multivariate Analysis of Variance Model Useful Especially for Growth Curve Problems," *Biometrika*, 51, 313–326.

Rao, C. R. (1965), "The Theory of Least Squares When the Parameters Are Stochastic and Its Application to the Analysis of Growth Curves," *Biometrika*, 52, 447–458.

——— (1967), "Least Squares Theory Using an Estimated Dispersion Matrix and Its Application to Measurement of Signals," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (Vol. 1), Berkeley: University of California Press, pp. 355–372.

——— (1973), *Linear Statistical Inference and Its Application* (2nd ed.), New York: John Wiley.

Reinsel, G. C. (1984), "Estimation and Prediction in a Multivariate Random Effects Generalized Linear Model," *Journal of the American Statistical Association*, 79, 406–414.

——— (1985), "Mean Squared Error Properties of Empirical Bayes Estimators in a Multivariate Random Effects General Linear Model," *Journal of the American Statistical Association*, 80, 642–650.

Reinsel, G. C., and Wincek, M. A. (1987), "Asymptotic Distribution of Parameter Estimators for Nonconsecutively Observed Time Series," *Biometrika*, 74, 115–124.

Silvey, S. D. (1959), "The Lagrangian Multiplier Test," *The Annals of Mathematical Statistics*, 30, 389–407.

Ware, J. H. (1985), "Linear Models for the Analysis of Longitudinal Studies," *The American Statistician*, 39, 95–101.

Wincek, M. A., and Reinsel, G. C. (1986), "An Exact Maximum Likelihood Estimation Procedure for Regression-ARMA Time Series Models With Possibly Nonconsecutive Data," *Journal of the Royal Statistical Society*, Ser. B, 48, 303–313.

Zerbe, G. O. (1979), "Randomization Analysis of the Completely Randomized Design Extended to Growth and Response Curves," *Journal of the American Statistical Association*, 74, 215–221.