121

6 MODELS FOR MULTILEVEL RESPONSE VARIABLES WITH AN APPLICATION TO GROWTH CURVES

Harvey Goldstein
Institute of Education
University of London

1 Introduction

In the majority of applications of multilevel models, one or more response variables are assumed to be measured at the lowest level (1) of the hierarchy. In general, however, we can model simultaneously sets of response variables measured at different levels. Thus, in a 2-level model for students grouped within classrooms, we could have achievement test scores measured on students together with an attitude measure for teachers or even the between-student standard deviation of test scores as a measure of class heterogeneity, where the variables at both levels are responses regarded as functions of, say, previous achievements and attitudes. Likewise, in the 2-level repeated measures model, which we consider in detail below, we have successive height measurements on individuals (at level 1) together with the individual's final adult height (at level 2). Simultaneous modeling of these response variables will allow us to estimate, among other things, the relationship between adult height and the parameters of pre-adult growth.

The use of height growth data provides a convenient illustration of the usefulness of the multilevel responses model. Much is already known about modeling such data, and systems for predicting adult height have been in use for some time, thus allowing a comparative evaluation. Nevertheless, the same models can be applied to educational data, and this is taken up in the discussion section.

Next we specify the model and illustrate it with a numerical example. The present paper describes the results of work in progress, and as will be evident from what follows, there still remains a number of important issues to be resolved.

2 The 2-Level Multivariate Model

As pointed out by Goldstein (1986a), a convenient characterization of the multivariate multilevel model is obtained by adding a further level below that of the lowest level actual unit, where dummy explanatory variables specify the response variables and the coefficients of these dummy variables are random at one or more higher levels. Thus, in the simple case of a single-level multivariate model with p possible measurements on each subject, we can write:

$$y_{ij} = \Sigma_s \beta_{is} x_{ijs} \qquad s = 1, \ldots, p \tag{1}$$
$$\beta_{ij} = \beta_j + v_{ij}$$
$$var(v_{ij}) = \sigma^2_{v,j}$$
$$cov(v_{ij}, v_{ik}) = \sigma_{v,jk}$$

where x_{ijs} is a (0,1) dummy variable which is one if $s = j$, and zero otherwise. The subscript j indexes a set of p different measurements on a subject. The β_j are the means of the measurements and the $\sigma^2_{v,j}$ and $\sigma_{v,jk}$ are simply the between-subject variances and covariances of the p measurements. If we add a set of further explanatory variables, z, at the subject level we obtain

$$y_{ij} = \Sigma_s \Sigma_m \alpha_{ms} z_{mij} x_{ijs} + \Sigma_s \beta_{is} x_{ijs} \tag{2}$$

This is one way of writing the multivariate linear model. We see that it allows each of the p measurements to have its own set of coefficients for the variable z, and that the subscript j, added to the z variables, allows measurement-specific explanatory variables. When, either by design or accident, not all the p response variables are measured on every subject, we can still obtain efficient estimates for the parameters in (2).

The basic 2-level repeated measures growth curve model can be written as:

$$\begin{aligned} y_{ij} &= \Sigma_m \alpha_m z_{mij} + \Sigma_t \beta_{tij} x^t_{ij} & t = 0, \ldots, q \\ \beta_{tij} &= \beta_t + u_{ti} + e_{tij} & t = 0, \ldots, q \end{aligned} \tag{3}$$

where x_{ij} is time or age. In the simplest case there is a single random variable at level 1 (occasions), namely e_{oij} representing a constant within-subject variance about the growth curve. We assume

$$cov(u_{ti}, e_{tij}) = cov(e_{tij}, e_{tik}) = 0 \ .$$

The independence of the level 1 residuals is a strong assumption in some applications. For example, growth in height has a seasonal component, so that where two or more measurements during a year are made this component will be superimposed on the underlying growth curve. Failure to model this effect will result in dependencies among the level 1 residuals. Although work on such models is currently being pursued, in the present paper we shall attempt to avoid this problem by using only yearly measurements on the subjects. Nevertheless, as we show below, we can still model, in a simple fashion, the level 1 variance as a function of age or other variables.

The u_{ti} are random variables at level 2 (subjects), giving rise to $q+1$ variances and $q(q+1)/2$ covariances. The z_{mij} are covariates which may vary from occasion to occasion. The coefficients α_m may be fixed or random. This model is discussed in detail, for example, by Strenio *et al.* (1983) and Goldstein (1986b). In general, the u_{ti} for large values of t will be set to zero, so that random between-subject variation is described by the low order polynomial coefficients.

Models (2) and (3) can be combined into a new class of very general models in which each of the response measurements has a separate polynomial regression on age with its own set of covariates, and where the random coefficients are correlated across the measurements at each level. Thus, for example, if height and weight are repeatedly measured in growing children, the intercept, linear etc. growth curve coefficients of height, at the subject level, will be correlated among themselves and with those for weight. An advantage of such a multivariate model is that, via these intercorrelations, it can provide efficient estimates for measurements with large numbers of randomly missing measurements. A simple bivariate example is given in Goldstein (1986b).

In the present paper we consider an example which is a bivariate specialization of the general model. The first variate is height, modeled as a function of age and certain covariates, and the second is adult height modeled simply as a function of the overall mean. There are two populations of subjects which have been sampled.

The model is written as follows:

$$\begin{aligned} y_{ij} &= \alpha_{1i}\delta_{ij}z_{1i} + \alpha_{2i}\delta_{ij}z_{2ij} + \delta_{ij}(\Sigma_t\beta_{tij}x_{ij}^t) \\ &+ (1-\delta_{ij})\gamma_i + (1-\delta_{ij})\alpha_{3i}z_{1i}\,, \qquad t = 0,\ldots,5\,, \qquad (4) \end{aligned}$$

where z_{1i} is a dummy (0, 1) variable indicating whether the subject belongs to group 1 or 2, and is thus a measurement made at the subject level. The variable z_{2ij} is the subject's bone age, estimated from a wrist radiograph according to Tanner *et al.* (1983). The variable δ_{ij} is 1 if the response is made during the growth period and 0 if adult height is measured, and x_{ij} is age, measured about a suitable origin. The coefficients α_{1i}, α_{2i}, and α_{3i}

are assumed to be fixed, and the remaining coefficients are assumed to be random as follows:

$$\begin{aligned}
\beta_{0ij} &= \beta_0 + u_{0i} + e_{0ij} & t = 0,1 \\
\beta_{1ij} &= \beta_1 + u_{1i} + e_{1ij} & \\
\beta_{2ij} &= \beta_2 + u_{2i} & \\
\beta_{3ij} &= \beta_3 + u_{3i} & \\
\beta_{tij} &= \beta_t & t = 1,\ldots,5 \\
\gamma_i &= \gamma_0 + v_i &
\end{aligned}$$

At the subject level, the random variables u_{0i}, u_{1i}, u_{2i}, and u_{3i}, v_i, have a 5-variate distribution with a zero mean vector and dispersion matrix Ω_2. At level 1, e_{0ij} and e_{1ij} have a bivariate distribution with a zero mean vector and dispersion matrix Ω_1. Thus, at any given age during the growth period, the variance of y_{ij} is given by:

$$X_2^T \Omega_2 X_2 + X_1^T \Omega_1 X_1 \ ,$$

where

$$\begin{aligned}
X_2^T &= (1, x_{ij}, x_{ij}^2, x_{ij}^3) \\
X_1^T &= (1, x_{ij}) \ .
\end{aligned}$$

The age range of growth considered in the example is 10 years to 18 years together with measurements of adult height in a sample of girls. During this period it is well known that there is a maximum of the velocity of growth at puberty and a minimum velocity approached as growth slows down at the approach to adulthood. There is also a pre-pubertal minimum of the velocity but for nearly all girls this occurs before the age of 10 years (Goldstein, 1986b). It is also well known that the ages of occurrence of these zero "acceleration" points vary between subjects.

To capture these growth features we require at least that growth coefficients up to the cubic vary randomly between subjects, since the age of zero acceleration is estimated by setting the second differential of the growth curve with respect to age, to zero. In fact, because the sample size in the example is relatively small (93 subjects), it has not been possible to fit random coefficients higher than the quadratic. A possible solution to this problem is to restrict the age range of the measurements so that the older ages which help to define the second stationary value of the velocity are excluded. A difficulty with this approach is in deciding where to truncate the age range. A further difficulty occurs where we wish to predict using values beyond the truncation point. We return to this in the example below.

3 Estimation and Prediction

The estimation procedure used is that described in Goldstein (1986a) and Goldstein (1987), namely iterative generalized least squares (IGLS), which is maximum likelihood when the random variables have a multivariate Gaussian distribution. Software written at the London Institute of Education was used (see Appendix).

Our interest is primarily in predicting γ_i, the adult height for individuals not in the sample. The mean γ_0 is obtained from the model estimates and we can form a posterior estimate of v_i in the usual manner. As is typically done when using such procedures, we ignore the sampling error of the random parameter estimates when calculating the standard errors of the predictions. In fact, the sample size of 93 cases appears to be large enough to justify this. Explicit formulae for the prediction equations and the standard errors of the predicted values are given in Goldstein (1987).

4 Data Analysis

The data for this example are measurements on two samples of girls measured from just after birth to adulthood. The first sample, known as the International Children's Centre London sample (ICC) consists of 52 girls born in the early 1950's in an area of central London. The second sample (NCH) consists of 41 girls in a children's home in Hertfordshire measured from entry to the home until adulthood. In both samples the children were measured close to their birthdays, and more frequently during periods of rapid growth. We have selected the yearly measurements from the 10th birthday onwards. Further details of the samples are given in Tanner *et al.* (1983). At each measuring occasion height was measured and bone age assessed according to the Tanner-Whitehouse scale (Tanner *et al.*, 1983).

Table 1 gives the parameter estimates from fitting the model (4). The term for study difference during growth was very small and has been omitted; the degree 5 polynomial term has been omitted for the same reason. Also, there is only a very small relationship between the level 1 variance and age, and so only a simple level 1 variance term has been fitted.

The ages of maximum height velocity are obtained by solving the following equation:

$$\beta_{2ij} + 3\beta_3 x + 6\beta_4 x^2 = 0 \qquad (5)$$

If we use the estimate for the variance of β_{2ij} in Table 1 and assume that β_{2ij} has a Gaussian distribution, then we can estimate the distribution of x.

TABLE 1

Height Related to Age, Bone Age, and Group: Girls Aged 10-18 Years

Fixed Coefficients	A Estimate	S.E.	B Estimate	S.E.
Adult height	162.0	0.59	162.1	0.59
Growth curve intpt	154.6	0.69	154.7	0.73
Bone age	0.59	0.11	–	–
Group (adult)	0.47	0.40	0.28	0.42
Age	4.68	0.17	5.20	0.14
Age^2	0.84	0.05	−0.87	0.06
Age^3	−0.142	0.008	−0.144	0.008
Age^4	−0.012	0.002	0.011	0.002

Random Coefficients

Model A:

Level 2	Covariance Matrix (Correlations) Adult Height	Growth Intpt	Age	Age^2
Adult height	29.2			
Growth intpt	27.4 (0.77)	43.7		
Age	0.70 (0.10)	−38.75(−0.40)	1.53	
Age^2	0.0084(0.00)	−1.69(−0.60)	0.45(0.86)	0.18

Level 1 variance = 0.63, s.e. = 0.06

Model B:

Level 2	Covariance Matrix (Correlations) Adult Height	Growth Intpt	Age	Age^2
Adult height	29.3			
Growth intpt	27.7 (0.74)	48.3		
Age	0.63(0.09)	−4.35(−0.50)	1.56	
Age^2	−0.01(0.00)	−1.96(−0.65)	0.46(0.84)	0.195

Level 1 variance = 0.61, s.e. = 0.06

Group is coded 1 if in ICC sample, 0 if in NCH sample.
Age is measured about an origin of 13.0 years.
Number of subjects = 93.
Number of measurements = 524.

This is done conveniently by using simulation, and Table 2 gives estimates for some percentiles of this distribution.

The mean age of 11.5 years is about 0.4 years lower than that found by Tanner *et al.* (1976) using a sample of the NCH children, including the measurements made every three months, by a method based upon smoothing each individual subject's height with a logistic curve.

As a further check on the model, standardized (shrunken) residuals at level 1 and level 2 are displayed in Figures 1-4.

As can be seen from Figure 1, the level 1 standardized residuals, plotted against the predicted height, are well behaved. Figures 2-4 are a selection of plots of the standardized level 2 residuals plotted against each other. It appears that the distribution of the slope residuals is negatively skewed, which may be a consequence of the absence of a random cubic coefficient.

Turning to the prediction of adult height, we first study the performance of the prediction on the 93 sample individuals, all of whom had an adult height measurement, by using the set of growth period measurements without adult height to predict the latter. These predicted heights are plotted against the actual adult heights in Figure 5, which shows no sign of any departure from linearity.

In this case we have used several repeated measures on each individual. In practice, however, a prediction is required from just one or two measurements. In Figure 6 we plot the predictions against adult height using the subsample of 62 girls who have a measurement within 0.1 years of their 12th birthday.

TABLE 2

Estimated Percentiles of the Age of Maximum Height Velocity

Percentile	Age
5	10.4
10	10.6
50	11.4
90	12.4
95	13.7

Mean age = 11.5 years

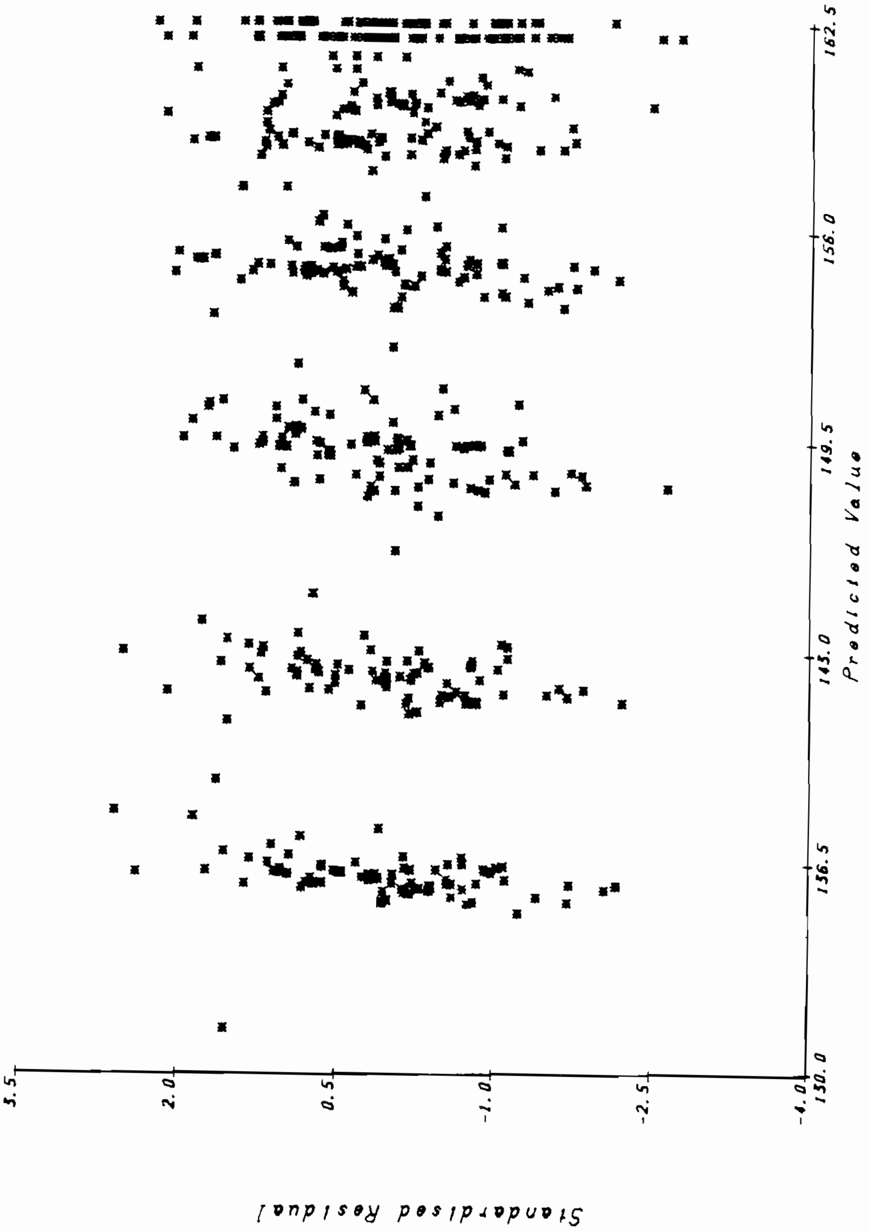

Fig. 1 Level 1 residuals.

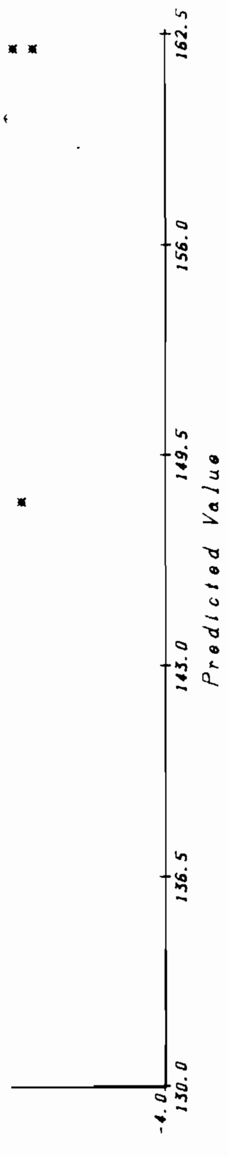

Fig. 1 Level 1 residuals.

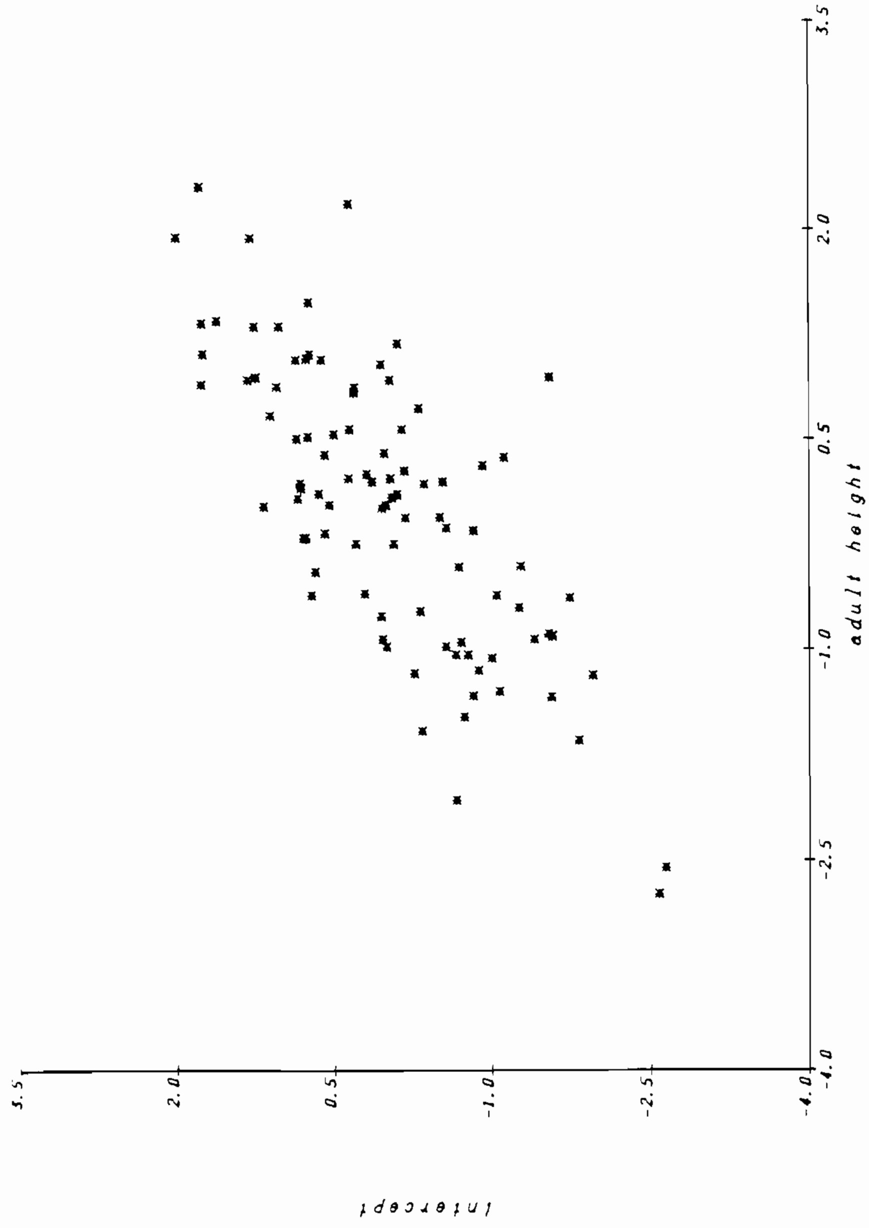

Fig. 2 Standardised residuals.

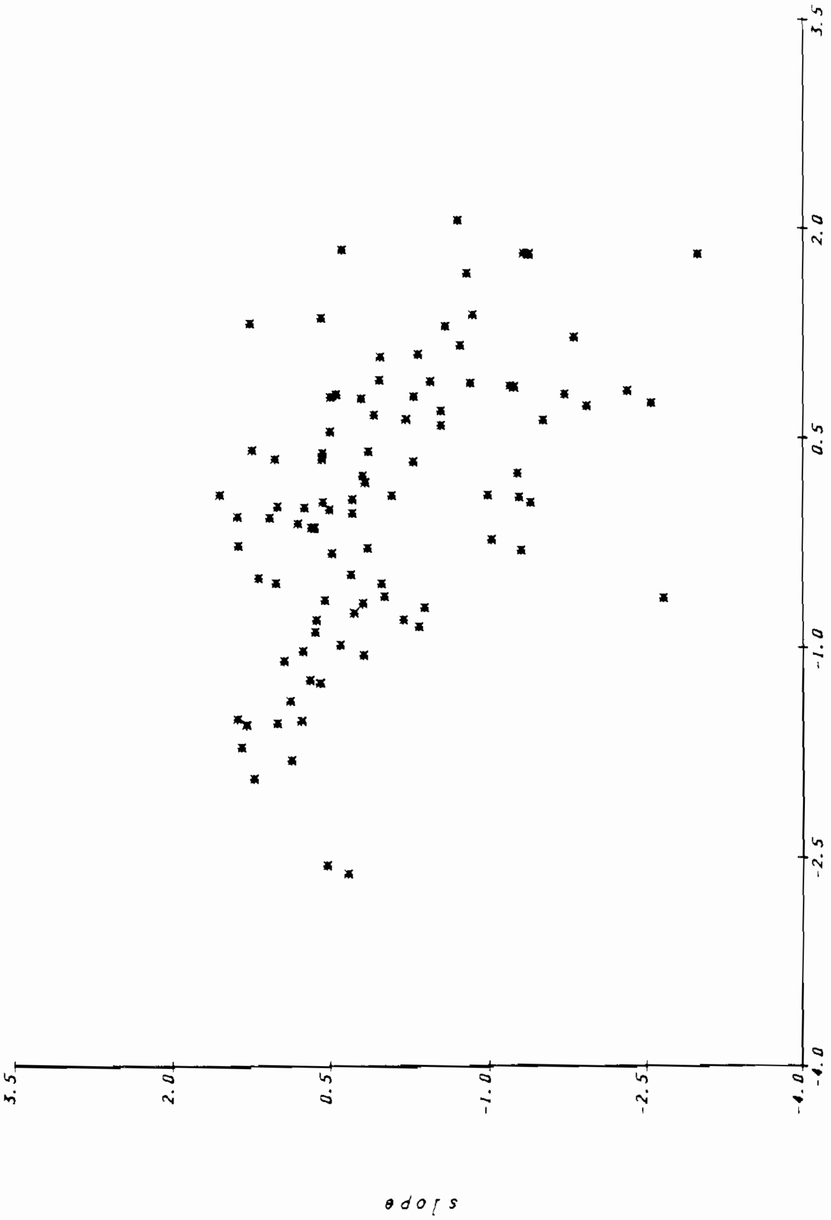

Fig. 3 Standardised residuals.

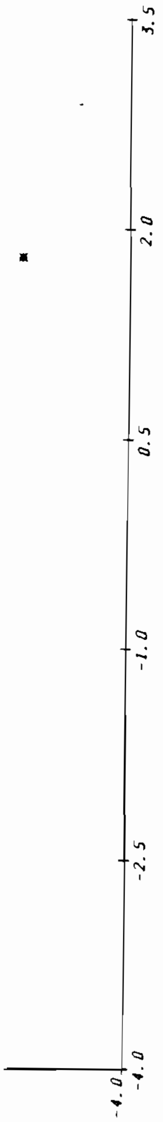

Fig. 3 Standardised residuals.

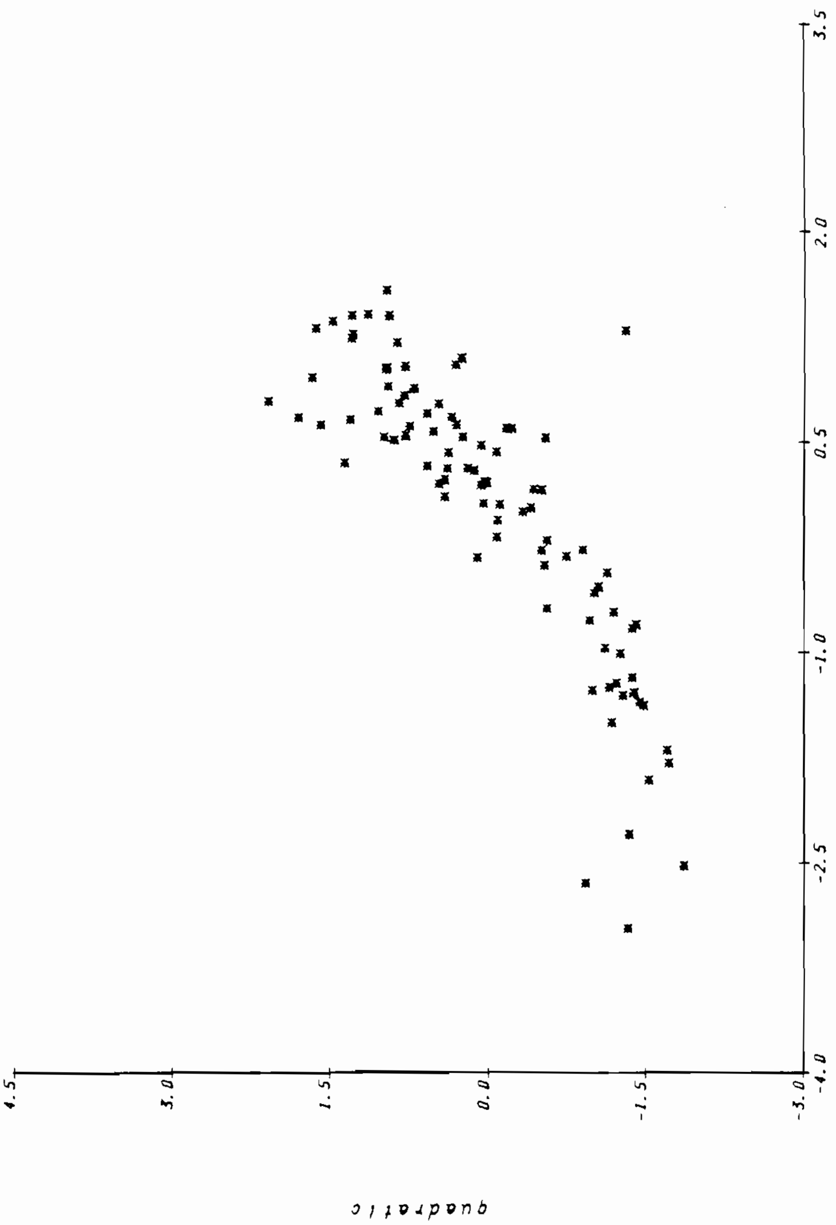

Fig. 4 Standardised residuals.

We see a linear relationship as before but with, as would be expected, a larger scatter of the observed residuals about their predictions. At this age the estimated standard error of the prediction is 3.7 cm. If two measurements are taken a year apart, centered at 12.0 years, the standard error becomes 3.0 cm. If six yearly measurements are used from 10 to 15 years, the standard error falls to 1.6 cm.

As already mentioned, one solution to the problem of being unable to fit a high enough order polynomial is to restrict the age range so that only one stationary value of the velocity occurs. We have therefore rerun the analyses excluding all ages greater than 15 years. The estimated parameters are shown in Table 3, and we obtain a mean age for the maximum height velocity of 11.5 years; however, it is possible to fit a random cubic coefficient in the model, and from that analysis we obtain a mean age of maximum height velocity of 11.7 years. It appears, however, that the individuals used in the present analysis (Professor J. M. Tanner, personal communication) may tend to have earlier-than-average pubertal growth. Children who left the study without adult height measures were not included in the present analysis and would tend to be late-maturing individuals, and this seems a plausible explanation for the relatively early ages of maximum height velocity in the present analyses.

5 Discussion

The analyses in this paper have demonstrated the feasibility of using a 2-level model for predicting adult height from serial measurements, taken during the growth period. Clearly, the method can be extended to other measurements, and we can also consider the multivariate case where several measurements are modeled jointly. In addition, the adult measurements to be predicted need not be those measured during growth, and this provides a flexible approach to the modeling of general repeated measures data. For routine use, a program can be written to make predictions with associated interval estimates, and the prediction can be updated as further measurements become available. It should be noted that the adult height predictions are generally population dependent. In the present analysis the group difference is small (0.5 cm.), but we cannot necessarily assume in general that all population differences will have been taken into account by conditioning on growth measurements. This will be a matter for empirical study. Likewise, it will often be necessary to adjust for a "secular trend" in adult height which has occurred between the time period when the data were collected and the period the results are in use.

. as would be expected. a
r predictions. At this age
cm. If two measurements
tandard error becomes 3.0
to 15 years, the standard

oblem of being unable to
he age range so that only
have therefore rerun the
The estimated parameters
for the maximum height
random cubic coefficient
mean age of maximum
that the individuals used
personal communication)
owth. Children who left
included in the present
iduals, and this seems a
maximum height velocity

asibility of using a 2-level
ements, taken during the
to other measurements,
re several measurements
rements to be predicted
is provides a flexible ap-
es data. For routine use,
associated interval esti-
r measurements become
predictions are generally
group difference is small
eral that all population
itioning on growth mea-
y. Likewise, it will often
ult height which has oc-
collected and the period

TABLE 3

Height Related to Age, Bone Age, and Group: Girls Aged 10-15 Years

	A		B	
Fixed Coefficients	Estimate	S.E.	Estimate	S.E.
Adult height	162.0	0.58	162.1	0.58
Growth curve intpt	154.7	0.69	154.8	0.72
Bone age	0.64	0.12	–	–
Group (adult)	0.56	0.33	0.28	0.33
Age	4.81	0.20	5.26	0.18
Age^2	−0.79	0.08	−0.87	0.08
Age^3	−0.19	0.03	−0.16	0.03
Age^4	−0.005	0.01	0.008	0.01

Random Coefficients

Model A:

Level 2 — Covariance Matrix(Correlations)

	Adult Height	Growth Intpt	Age	Age^2
Adult height	29.2			
Growth intpt	27.7 (0.78)	42.9		
Age	0.96(0.13)	−4.20(−0.46)	1.95	
Age^2	0.02(0.01)	−1.92(−0.60)	0.59(0.86)	0.24

Level 1 variance = 0.49, s.e. = 0.06

Model B:

Level 2 — Covariance Matrix(Correlations)

	Adult Height	Growth Intpt	Age	Age^2
Adult height	29.3			
Growth intpt	27.7 (0.78)	42.9		
Age	1.12(0.15)	−4.70(-0.51)	1.96	
Age^2	0.11(0.04)	−2.12(-0.65)	0.60(0.86)	0.25

Level 1 variance = 0.50, s.e. = 0.06

Group is coded 1 if in ICC sample, 0 if in NCH sample.
Age is measured about an origin of 13.0 years.
Number of subjects = 93.
Number of measurements = 446.

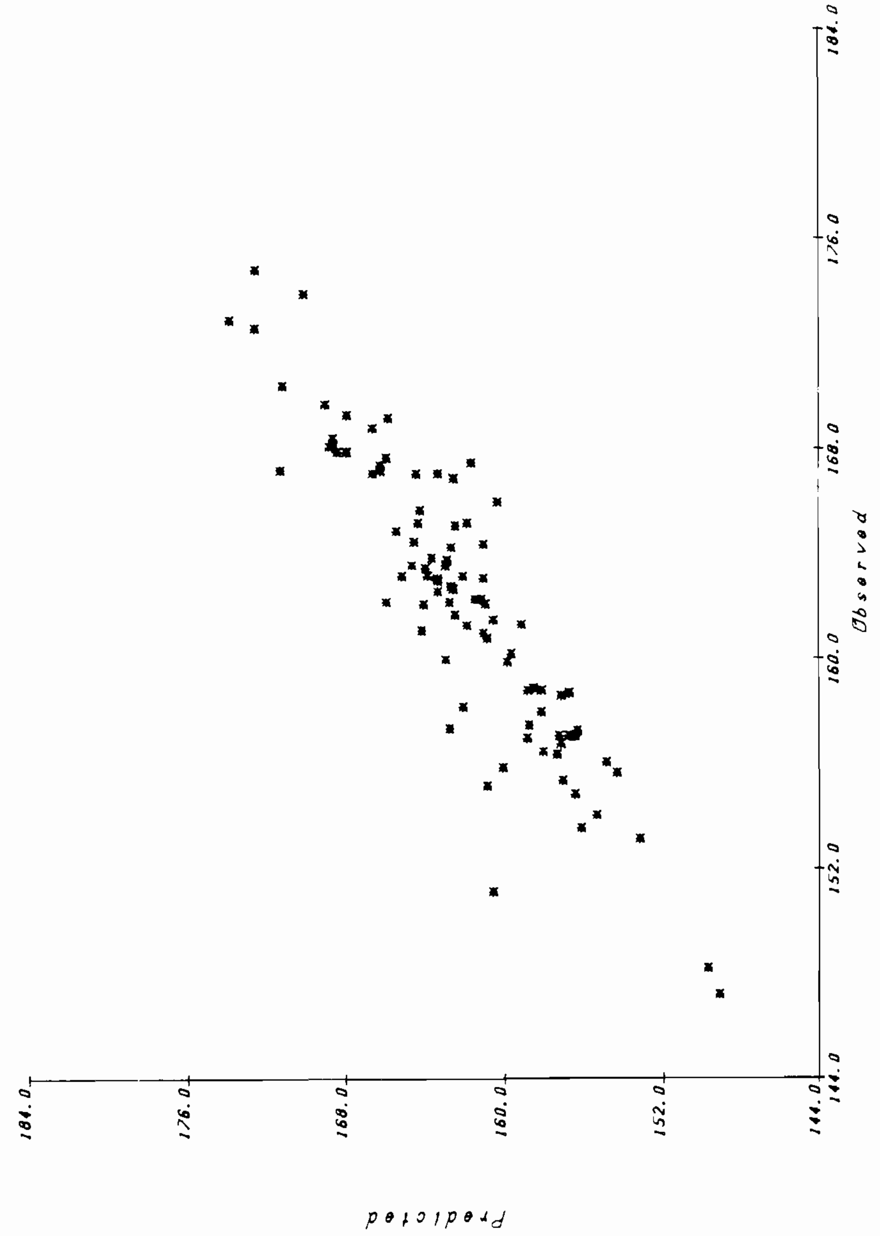

Fig. 5 Girls adult height prediction.

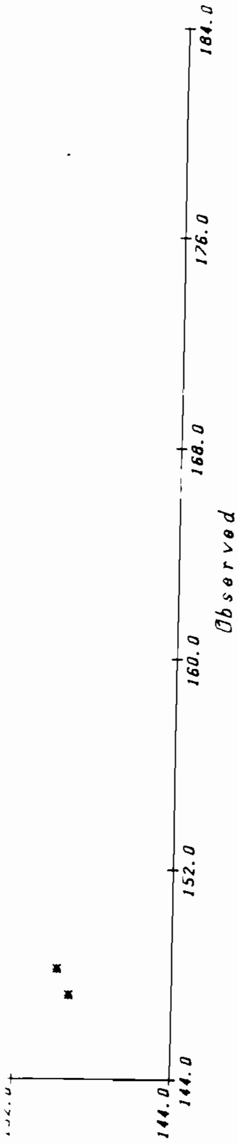

Fig. 5 Girls adult height prediction.

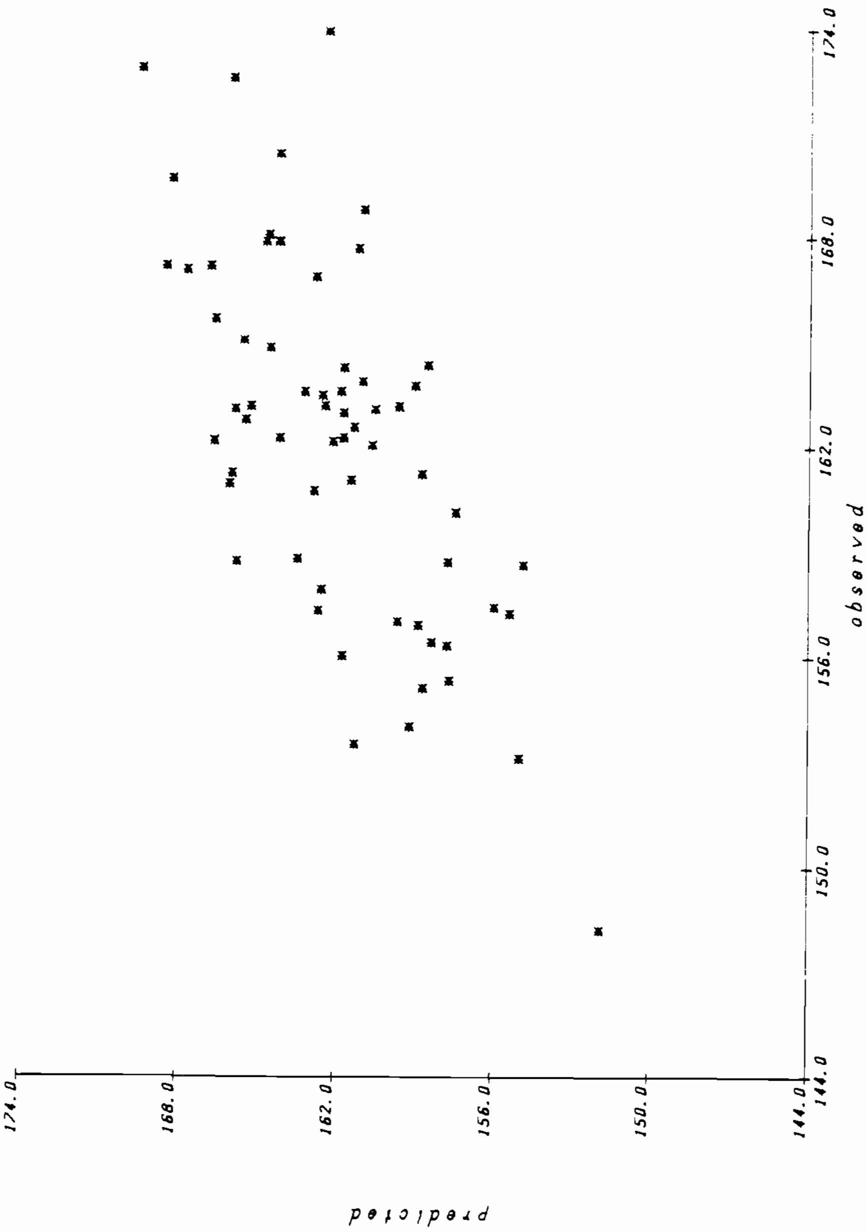

Fig. 6 Girls age 12 adult height prediction.

The models in this paper can be used with longitudinal educational data where interest lies in the prediction of an outcome such as an examination score. In this case, serial measurements may be age- and population-standardized, so that the means of the growth coefficients are zero. Where different serial measurements are made, these can be treated as separate response variables with coefficients correlated at level 2. Multiple outcome variables can be handled by direct extension. In such models, covariates and group differences are often of interest in their own right. Furthermore, a third level, that of the school, will generally be needed and the model coefficients in general may vary across schools. In this case, where predictions are required for individual students, estimates of school level (shrunken) residuals should be incorporated where the student's school is known.

We have assumed simple multivariate distributions among the measurements and the random parameters. In fact, in the case of height data there are some constraints which ought to be included in the model, namely that, for any individual, the adult height cannot be less than any of the growth measurements. Thus, using the models in this paper it would be possible to predict an adult measurement less than the most recent growth measurement. This especially will be the case for growth measures taken towards the end of the growth period, and provides another reason for restricting prediction to ages up to 15 years only. The problem is one which affects all height prediction methods and needs further study.

Two other methods are in use for prediction of adult height. The one by Tanner *et al.* (1983) is based upon separate regression predictions of height at each age, or pair of ages. While this procedure can in principle produce efficient predictions, it is not very flexible. Thus, the accuracy of the prediction equation is limited by the actual number of subjects at the age being used, whereas the 2-level model procedure can use efficiently all the data available, including those cases without an adult height measurement. Also the fixed-age prediction method cannot realistically handle more than two serial measures, whereas the 2-level procedure can include as many as are available.

The other procedure (Bock, 1986) is similar to the present one but instead uses a non-linear model fitted to the whole growth age range with parameters varying between subjects.

Detailed comparisons of these procedures have not yet been carried out. The 2-level polynomial model, however, would seem to be the most flexible and potentially the most efficient of these methods. It can easily handle multiple measurements, it can model within individual changes in variation, it can make use of data from individuals with only very few measurements, and it can handle measurements other than height, for which simple non-linear growth models are unavailable.

A further development would be to extend the number of covariates in the model. Thus, Tanner *et al.* (1983) effectively include the occurrence of menarche as a covariate by presenting separate predictions for those girls who have and who have not yet experienced that event. Likewise, other stages of pubertal development could be included. The inclusion of subject-level variables such as parental height and birth order might also be useful. In some cases, it may be preferable to treat a continuous, occasion-related covariate as a response. Thus, we could fit a bivariate growth model to height and bone age, where in a simple model bone age might be a quadratic function of age with all the coefficients random at the subject level, the intercept and quadratic coefficients having a mean value of zero and the linear coefficient having a mean value of 1.0. The predictor of adult height would then be a function of the set of height and bone age residuals. An important advantage of this model is that even where bone age is not measured at all occasions, all the available bone age measurements can be used in the prediction. This contrasts with the present models where we use either all the occasions without bone age or just those that contain bone age.

Finally, it should be stressed that large samples are important to secure stable estimates and to enable higher-order fixed and random coefficients to be included so that the model can be properly specified. It would be convenient, for example, to be able to model a much wider age range than that considered here, in order to avoid the problem discussed at the end of Section 4. That would require further higher-order random coefficients to cope with at least two more stationary values of height growth in the pre-pubertal period. The optimum combination of overlapping age ranges is a matter for further empirical study. Further work is also needed on the modeling of measurements made close together in time where serial correlations will be present at level 1.

6 APPENDIX

MULTILEVEL: A Model-fitting Program using Iterative Generalized Least Squares

This brief overview describes the data input structure, the facilities for specifying a model, and the form of output. Running specifications can be obtained from the author and a copy of the current version, written in FORTRAN 77, is available.

6.1 Data Structure

The basic data record is a level-1 unit. These are nested within each level-2 unit and level-2 units are nested within level-3 units. The sequence of data records thus corresponds to the hierarchical structure. The facilities for data transformation within the program currently are somewhat limited. It is, however, possible to form the square of any explanatory variable, compute the mean value of any explanatory variable for all the level-2 or level-3 units to form new aggregated variables, and to calculate the standard deviation of a level-1 explanatory variable for each level-2 unit. In addition, one can specify whether or not an explanatory variable is to be measured from its overall mean.

6.2 Model Specification

The fixed and random parts of the model are specified separately. For the fixed part, any selection of the input explanatory variables can be made. For the random part, the parameters are specified separately at each level. At each level there are potentially $n(n+1)/2$ variances and covariances, where n is the number of input explanatory variables (including the constant term if required) and thus also the number of potential random coefficients. Any coefficient can be defined as random at any level. There is also a facility for allowing a non-zero covariance to be estimated when one of the corresponding variances is zero (see Goldstein, 1987, for a justification of this).

The program also allows for the specification of no random variables at level 1 in order to fit multivariate models. Starting values can be input, with the default being OLS values.

6.3 Output

Output is flexible. It allows printing of estimated covariance matrices of estimated parameters during iterations. It allows printing of raw data. It will calculate and print shrunken level-2 and level-3 residuals together with conditional and unconditional standard error estimates. Iterations are controlled by a relative accuracy convergence criterion on the random parameters.

6.4 General

The program is written in FORTRAN 77, to be as general as possible and handle a wide class of standard models including loglinear discrete response multilevel models (Goldstein, 1987). The raw data need to be stored for use at each iteration in order to allow flexibility in specifying the level-1 covariance structure. The standard version of the current program allows 2500 level-1 records, 120 level-2 and 120 level-3 units, up to 55 random parameters at each level and up to 20 explanatory variables. The major deficiency lies in its heavy demand on CPU time. A new version which will be available in 1988 will be both faster and more economical in use of storage.

REFERENCES

Bock, R. D. (1986). Unusual growth patterns in the Fels data. In *Human Growth: A Multidisciplinary Review* (Ed. by A. Demirjian). London and Philadelphia: Taylor and Francis.

Goldstein, H. (1986a). Multilevel mixed linear model analysis using iterative generalised least squares. *Biometrika*, **73**, 43-56.

Goldstein, H. (1986b). Efficient statistical modelling of longitudinal data. *Annals of Human Biology*, **13**, 129-141.

Goldstein, H. (1987). *Multilevel models in educational and social research.* London: Griffin; New York: Oxford University Press.

Strenio, J., Weisberg, H. I., & Bryk, A. S. (1983). Empirical Bayes estimation of individual growth curve parameters and their relationship to covariates. *Biometrics*, **39**, 71-86.

Tanner, J. M., Whitehouse, R. H., Cameron, N., Marshall, W. A., Healy, M. J. R., & Goldstein, H. (1983). *Assessment of Skeletal Maturity and Prediction of Adult Height (TW2 Method).* London: Academic Press.

Tanner, J. M., Whitehouse, R. H., Marubini, E., & Resele, L. F. (1976). The adolescent growth spurt of boys and girls of the Harpenden Growth Study. *Annals of Human Biology*, **3**, 109-126.