

A MIXED-EFFECTS MODEL FOR GROWTH CURVES
ANALYSIS IN A TWO-WAY CROSSED
CLASSIFICATION LAYOUT

MARIO MIGUEL OJEDA* HARDEO SAHAI†

Received/Recibido: 11 Nov 2003

Abstract

We propose a mixed-effects linear model for analyzing growth curves data obtained using a two-way classification experiment. The model combines an unconstrained means model and a regression model on the time, in which the coefficients are considered random. The model allows for experimental unit covariates so as to study the trend and the variability of the individual growth curves. Comments on data analysis strategies are provided. An application of the model is illustrated using a data-set comes from a chrysanthemum growth experiment.

Keywords: multilevel linear regression models, random coefficients models, means models, data analysis strategies.

Resumen

Proponemos un modelo lineal de efectos mixtos para analizar datos de curvas de crecimiento de un experimento con dos criterios de clasificación. El modelo combina un modelo no restringido de medias y un modelo de regresión sobre el tiempo, en el cual los coeficientes son considerados aleatorios. El modelo considera covariables a nivel de la unidad experimental para estudiar la tendencia y la variabilidad de las curvas de crecimiento. Se proporcionan comentarios sobre estrategias de análisis de datos. Se ilustra la aplicación del modelo usando un conjunto de datos de un experimento de crecimiento de crisantemos.

Palabras clave: Modelos de regresión lineal multinivel, modelos de coeficientes aleatorios, modelos de medias, estrategias de análisis de datos.

Mathematics Subject Classification: 62J99, 62P10.

*Facultad de Estadística, Universidad Veracruzana, Apartado Postal 475, Xalapa Veracruz, México.
E-Mail: mojeda@uv.mx

†Department of Biostatistics and Epidemiology, University of Puerto Rico, P.O. Box 365067, San Juan, Puerto Rico 00936. E-Mail: hardeosahai@yahoo.com

1 Introduction

The use of linear complex models, which include both fixed and random effects and hierarchically nested dataset, are becoming increasingly common in many areas of applied research; for example, in agricultural experimentation, industrial productivity improvement, behavioral and educational research, among others (Rawlings et al., 1998; Chapter 18; Gibbons, 2000; Gibbons and Bhanmik, 2001). Usually the models for repeated measures or growth curves consider a complex structure of variance-covariance matrix (Jenrich and Schulchter, 1986; Rencher, 2000). Measures at different points for each experimental unit can be considered as nested within the unit. Covariate data, both for time period and experimental unit, add additional complexities to the modelling process. Large scale agricultural experimentation, considers models which include environmental influences and repeated measures (Vonesh and Chinchilli, 1997).

For a planned experiment, where the experimenter is interested in studying the growth curves, the most common approach uses the fixed multivariate growth curve model (Potthoff and Rao, 1964; Seber, 1984). However, this approach does not consider the random variation between individual growth curves, which is a very important aspect of the problem. Following the contribution of Rao (1965), a random coefficient model approach has been proposed, in order to incorporate more real complexities in the modelling process (Laird and Ware, 1982; Ware, 1985). Presently, using a theoretically integrated approach, the multilevel linear models methodology (Goldstein, 1987; 1995) permits us to solve repeated measures and o growth curves data analysis problem in a more realistic framework. Multilevel linear models, also referred to as random coefficient models (Longford, 1993; 1995), covariance components models (Pienerson, 1986), mixed linear models (McClellan et al., 1991; Verbeke and Molenberghs, 1997; Gibbons, 2000), and hierarchical linear models (Bryk and Raudenbush, 1992), are currently in widespread use.

For growth curves modelling, several authors have recently considered some particular versions of this family of models (see, e.g., Hui, 1988; Goldstein, 1989a; Ojeda and Juárez-Cerrillo, 1996). Estimation principles and methods, hypothesis testing procedures, diagnostics tools, and some model checking algorithms, are now available for this kind of models (see Hand and Crowder, 1996 for a review on these topics). Moreover, various computational software (see Kreft et al., 1994 and Singer 1998) permit to implement strategies of data analysis using such models.

In this paper, we propose a two-level regression model for analyzing a two-way experiment with growth curves data. The model is a particular case of a mixed-effects linear model that includes an unconstrained means model (Hocking, 1985) and a regression model with random coefficients. Some comments on estimation, hypothesis testing and diagnostic checks for the modelling process are included. An application of the model is illustrated using a data-set comes from a fertilization study on *Chrysanthemum* plants.

2 The model

For denoting the individual growth curves data, let $\mathbf{y}_{ijk} = (y_{ijk1}, y_{ijk2}, \dots, y_{ijkm})'$, where y_{ijk} is the measurement on the k -th experimental unit; $k = 1, 2, \dots, n_{ij}$ in the (i, j) -th cell;

$i = 1, 2, \dots, a; j = 1, 2, \dots, b$, in a two-way classification, for the time $t; t = l, 2, \dots, m$. Let $\mathbf{w}_{ijk} = (w_{ijk1}, w_{ijk2}, \dots, w_{ijkq})'$ a vector of corresponding observations on W_1, W_2, \dots, W_q , the explanatory curve (level-two) variables. Then the growth profile for each cell is $\mu_{ij} = (\mu_{ij1}, \mu_{ij2}, \dots, \mu_{ijm})'$. The general form of the proposed model is:

$$y_{ijkt} = \mu_{ij} + \beta_{0s} + \sum_{l=1}^L x_{lijkt} \beta_{ls} + e_{ijkt}, \quad (1)$$

where $x_{lijkt}; l = 1, 2, \dots, L$, are the values on X_1, X_2, \dots, X_L explanatory time dependent (level-one) variables; $s = 1, 2, \dots, S$, with $S = \sum_{i=1}^a \sum_{j=1}^b n_{ij}$; $\mu_{ij} = \mu + \alpha_{A_i} + \alpha_{B_j} + \alpha_{AB(i,j)}$ and

$$\beta_{ls} = \gamma_{l0} + \sum_{r=1}^q w_{ijkr} \gamma_{lr} + u_{ls}; l = 0, 1, \dots, L; s = 1, 2, \dots, S, \quad (2)$$

where μ is the general mean, and α_{A_i} , α_{B_j} and $\alpha_{AB(i,j)}$ are the fixed-effects associated with the two-way classification.

The first part in model equation (1) is an unconstrained means model for the unbalanced two-way classification (see, e.g., Hocking, 1985; chapter 6), and the second part is a regression model that considers the time or time functions as explanatory variables; where the coefficients β_{ls} are associated with the growth trajectory for each experimental unit. The set of models in (2) constitutes a $(L + 1)$ variate regression model, where the response vectors are $\beta_s = (\beta_{0s}, \beta_{1s}, \dots, \beta_{Ls})'$, and the vectors $\mathbf{u}_s = (u_{0s}, u_{1s}, \dots, u_{Ls})'$ are the level-two random errors. If we let $\mathbf{w}_{ijk} = (1, w_{ijk1}, w_{ijk2}, \dots, w_{ijkq})'$, then we can write the level-two equation (2) in a compact matrix notation as

$$\beta_s = \mathbf{W}_s \mathbf{\Gamma} + \mathbf{u}_s; s = 1, 2, \dots, S, \quad (3)$$

where $\mathbf{W}_s = \mathbf{I}_{L+1} \otimes \mathbf{w}'_{ijk}$ with \otimes denoting the Kronecker product, \mathbf{I}_{L+1} is the $L + 1$ identity matrix, and $\mathbf{\Gamma} = (\gamma'_0, \gamma'_1, \dots, \gamma'_L)'$ is formed by horizontal stacking of $\gamma_l = (\gamma'_{l1}, \gamma'_{l2}, \dots, \gamma'_{lq})'$.

In order to write equation (1) in a compact matrix notation, let $\mathbf{x}_{ijkt} = (1, x_{lijt}, x_{2ijkt}, \dots, x_{Lijkt})'$ and $\Theta = (\mu, \alpha_{A_1}, \alpha_{A_2}, \dots, \alpha_{A_{(a-1)}}, \alpha_{B_1}, \alpha_{B_2}, \dots, \alpha_{B_{(b-1)}}, \alpha_{AB(1,1)}, \dots, \alpha_{AB[(a-1)(b-1)]})'$, and define

$$\begin{aligned} \alpha_{A_a} &= 1 - \sum_{i=1}^{a-1} \alpha_{A_i} \\ \alpha_{B_b} &= 1 - \sum_{j=1}^{b-1} \alpha_{B_j} \\ \alpha_{AB(i,b)} &= 1 - \sum_{j=1}^{b-1} \alpha_{AB(i,j)}; i = 1, 2, \dots, (a-1) \\ \alpha_{AB(a,j)} &= 1 - \sum_{i=1}^{a-1} \alpha_{AB(i,j)}; j = 1, 2, \dots, b. \end{aligned}$$

Then the model equation (1) can be written as

$$\mathbf{y}_{ijt} = \boldsymbol{\Theta}' \mathbf{Z}_{ij} + \mathbf{x}'_{ijt} \beta_s + \mathbf{e}_{ijt}, \quad (4)$$

where $\mathbf{y}_{ijt} = (y_{ij1t}, y_{ij2t}, \dots, y_{ijn_{ij}t})'$ is the outcome vector for the n_{ij} experimental units in the (i, j) -th cell for the fixed time t , \mathbf{Z}_{ij} is the design matrix for the two-way unconstrained means model (see, e.g., Hocking, 1985; pp. 137–139) and $\mathbf{e}_{ijt} = (e_{ij1t}, e_{ij2t}, \dots, e_{ijn_{ij}t})'$ is the corresponding vector of the one-level random errors.

The model in (3) and (4) is accompanied with the following general assumptions: $E\{\mathbf{e}_{ijt}\} = \mathbf{0}$, $Var\{\mathbf{e}_{ijt}\} = \Sigma$, $E\{\mathbf{u}_s\} = \mathbf{0}$, $Var\{\mathbf{u}_s\} = \boldsymbol{\Omega}$, and all the random vectors are uncorrelated, that is, $Cov\{\mathbf{e}_{ijt}, \mathbf{e}_{i^*j^*t^*}\} = 0$, $Cov\{\mathbf{u}_s, \mathbf{u}_{s^*}\} = 0$, $Cov\{\mathbf{u}_s, \mathbf{e}_{ijt}\} = 0$. The simpler form of this model arises when $\sum_i = \sigma^2 \mathbf{I}_m$, which is discussed by Laird and Ware (1992), Ojeda and Juárez-Cerrillo (1996), among others.

Substituting equation (3) in equation (4) and letting $\mathbf{Z}_s = \mathbf{T}\mathbf{W}_s$, where $\mathbf{T} = (\mathbf{J}_m \otimes \mathbf{x}_{ij})$ with \mathbf{J}_m indicating the vector of m ones, we can write the model in (4) as

$$\mathbf{y}_{ijt} = \boldsymbol{\Theta}' \mathbf{Z}_{ij} + \mathbf{Z}_s \boldsymbol{\Gamma} + \mathbf{T} \mathbf{u}_s + \mathbf{e}_{ijt}. \quad (5)$$

This is a particular case of the general mixed-effects model, which is again a special case of the general linear model

$$\mathbf{Y} = \mathbf{X}^* \boldsymbol{\beta} + \mathbf{e}^*, \quad (6)$$

where the variance-covariance matrix of the observation is expressed in general by $\mathbf{V} = Var\{\mathbf{e}^*\}$. The matrix \mathbf{X}^* may contain a large number of variables and its general form is rather very complex and we only use this notation for a simple discussion of the methods of estimation in this model.

If the variance-covariance matrix \mathbf{V} is known, we can apply the generalized least squares method and obtain the estimator:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^{*\prime} \mathbf{V}^{-1} \mathbf{X}^*)^{-1} \mathbf{X}^{*\prime} \mathbf{V}^{-1} \mathbf{Y}. \quad (7)$$

Note that the determination of $\hat{\boldsymbol{\beta}}$ may involve substantial computing effort and for a matrix \mathbf{V} with a complex structure a closed-form expressions for $\hat{\boldsymbol{\beta}}$ is not readily obtainable. In general, it is convenient to assume a simple structure of \mathbf{V} which can be expressed in terms of a few parameters. Often \mathbf{V} is unknown, and no information about the structure of \mathbf{V} is available. In the past, only some very simple structures of \mathbf{V} have been considered and no satisfactory solutions have been found. For a general review of this problem, see Henderson (1986). It is important to recognize that when hierarchical data have a more complex structure, as is typical in many real-life problems, iterative algorithms have to be used to compute the estimates of the fixed- and random-effect parameters.

In recent years, different approaches and algorithms have been developed for the evaluation of $\hat{\boldsymbol{\beta}}$ in (6). Under the assumptions of normality, three conceptually different approaches are: empirical Bayes estimation (Dempster et al., 1977); the restricted maximum likelihood (Mason et al., 1983), and the full maximum likelihood (Longford, 1987). In the empirical Bayes estimation. Dempster et al. (1981) applied the EM algorithm to complex hierarchical data; see also Laird and Ware (1982) and Strenio et al. (1983) for other

applications of this approach. Goldstein (1986) proposed the iteratively reweighted generalized least squares (IGLS) procedure and showed that the estimators thus obtained are equivalent to the full maximum likelihood estimators under the assumption of normality. Goldstein (1989b) adapted the IRGLS algorithm for the restricted maximum likelihood solution. Longford (1987), using a full maximum likelihood approach, considered the Fisher scoring algorithm, which was originally proposed by Hartley and Rao (1967); see also Longford (1993, pp. 24). All these algorithms have been implemented in a variety of software packages. For a comparative review of various software to implement these algorithms, see Kreft et al. (1994). Complex structures of the matrix \mathbf{V} can be generally considered; see, for example, Rasbash et al. (1996), Yang et al. (1996) and Verbeke and Molenberghs (1997).

On the other hand, tests of hypotheses in model (5) are formulated for the fixed effects, subject-level random coefficients, and for the variance and covariance components. For the fixed effects coefficients, the approach of hypothesis testing is based on the general linear hypothesis formulation; that is, the hypotheses that are testable in hierarchical linear models are: $H_0 : \mathbf{C} \times \text{vec}(\Psi) = \mathbf{0}$ vs. $H_1 : \mathbf{C} \times \text{vec}(\Psi) \neq \mathbf{0}$, where vec is the vectorization operator which stacks the columns of a matrix in a column vector, and $H_0 : \mathbf{C}\beta = \mathbf{0}$ vs. $H_1 : \mathbf{C}\beta \neq \mathbf{0}$. Under the assumptions of normality, these test procedures, as special cases of the general linear hypothesis, are exact: Hypothesis testing for variances and covariances is performed as $H_0 : \Omega = \Omega_0$ vs. $H_1 : \Omega \neq \Omega_0$, and is generally based on the likelihood-ratio statistic. Good reviews and bibliographical notes about the estimation problem and the hypothesis testing procedures in hierarchical linear models are now available in Bryk and Raudenbush (1992), Longford (1993), Goldstein (1995), Verbeke and Molenberghs (1997), Khuri et al., (1998), Rencher (2000) and Gibbons (2001).

3 Strategies for data analysis of growth curves in a two-way classification

Misconceptions and misuses in fitting complex models are very frequent in all areas of application, but specifically in growth curves analysis. The model proposed here, for an adequate use in a particular modeling process, requires an integrated data analysis strategy. Although specially tailored software, some general methodological guidelines, and application examples are currently available (see, e.g., Lindsey, 1993; Diggle et al., 1994; Hand and Crowder, 1996), frequent problems in the analysis of data are encountered and appropriate diagnostic tools to solve some of these difficulties are not readily available. An adequate model formulation and goodness of fit evaluation (diagnostic and sensitivity analysis) require to establish a series of steps, which must begin with an intensive exploratory data analysis. Plots and individual fits for all experimental units in each cell in the two-way classification model will produce a general idea of the growth trend variation between treatments. For each time period we could obtain a fit of the two-way unconstrained means model, producing and analyzing residuals. A sample profile analysis would produce a better preliminary view about the significance of the fixed-effects. All the evidence in these preliminary and exploratory analyses would generally provide an

adequate model postulation.

In order to evaluate the explanatory variables in the two-level model, a simple random coefficient regression model fit, including only the most important variables, is suggested as follows. Diagnostic checks and level-one residual analysis are carried out for evaluating the goodness of fit. If an explanatory variable is added each time, then the diagnostic checks of residuals is repeated once more. Such an iterative scheme allows us to increase the complexity of the model in a gradual manner. When the more parsimonious model being fitted is obtained, we need to implement a complete diagnostics check once again. Level-one and level-two residuals would be checked and evaluated in a more careful way. For level-one residuals, traditional diagnostics can be implemented, as checks for outlier identification, symmetry, normality, and the assumption of constant variance. Also, plots against explanatory variables could be used to detect additional variables to be included in the model. Some comments and suggestions about level-one residual analysis are included in Goldstein (1987, 1995), Bryk and Raudenbush (1992), Longford (1993) and Langford and Lewis (1998). Ojeda and Juárez-Cerrillo (1996) present guidelines and exploratory multivariate techniques for diagnostic checks of level-two residuals. Shi and Ojeda (2004) studied influence in regression models for growth curves. Predicted trend for each individual can be obtained in order to study the predicted variability patterns, by using multivariate techniques such as principal components analysis or cluster analysis (see, e.g., Ramsay and Silverman, 1997).

4 An illustrative example

We use growth data from an experiment conducted at the university of Veracruz, in Xalapa, Veracruz, México, whose objective was to study the effect of fertilization strategies on chrysanthemum plant growth. Details about this study were reported by Fussillier (1996), who presented a preliminary analysis of these data conducted using a two-way crossed classification balanced ($n_{ij} = 4$) experiment. The experimental factors were: “fertilization formulae”, with 2 levels ($a = 2$); and “base composition soil”, with 5 levels ($b = 5$). For each plant in the study, height (Y) was recorded during 15 ($m = 15$) consecutive weeks (see Figure 1). Following a general strategy for statistical modeling, the proposed ANOVA model for testing the significance of factor effects is

$$\mu_{ij} = \mu + \alpha_{A_i} + \alpha_{B_j} + \alpha_{AB(i,j)}, \quad (8)$$

where μ_{ij} represents the fixed-effect term due to cell-mean in an univariate ANOVA model for each week. The results of the ANOVA F tests including the p -values are displayed in Figure 2.

Using the evidence obtained from an exploratory data analysis, Hoerl’s linearized curves (Daniel and Wood, 1980, pp. 20–24) were evaluated for fitting the effect of response shapes on the time. In order to model the growth curve for each experimental unit, we considered the linear equation:

$$\ln(Y) = \ln(\beta_0) + \beta_1 \ln(t) + \beta_2 t, \quad (9)$$

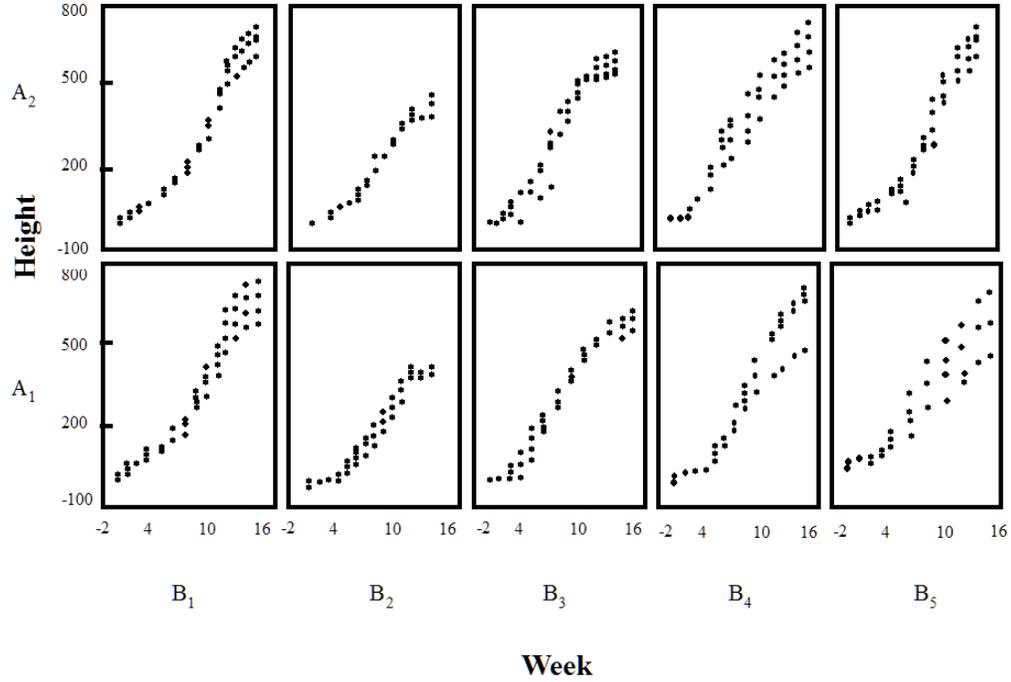


Figure 1: Categorized two-way crossed height trajectories for 15 consecutive weeks in a growth of chrysanthemum plants experiments.

where t represents the corresponding time in weeks. Then the proposed one-level growth model is:

$$y_{ijkt}^* = \mu_{ij} + \beta_{0s}^* + \beta_{1s}x_{1ijkt} + \beta_{2s}x_{2ijkt} + e_{ijkt}, \quad (10)$$

where $x_{1ijkt} = \ln(t)$, $x_{2ijkt} = t$ and $y_{ijkt}^* = \ln(y_{ijkt})$. A simple mean model was proposed for the random coefficients; i.e.,

$$\beta_{1s} = \gamma_l + u_{ls}; l = 0, 1, 2; s = 1, 2, \dots, 40. \quad (11)$$

The fitted model was obtained as:

$$\hat{y}_{ijkt}^* = \hat{\mu}_{ij} + 3.685_{(0.035)} + 0.272_{(0.032)}x_{1ijkt} + 0.151_{(0.006)}x_{2ijkt}, \quad (12)$$

where the standard errors are indicated within parenthesis just below the fitted values of the respective parameters. Both factor effects A and B and their interaction effect AB were highly significant. The unique significant variability in the regression coefficients was in intercepts ($\hat{\sigma}_0^2 = 0.0038_{(0.0014)}$), which represents 8.3% of the total error variance. Table 1 presents the fixed-effect components for the cell means $\hat{\mu}_{ij}$.

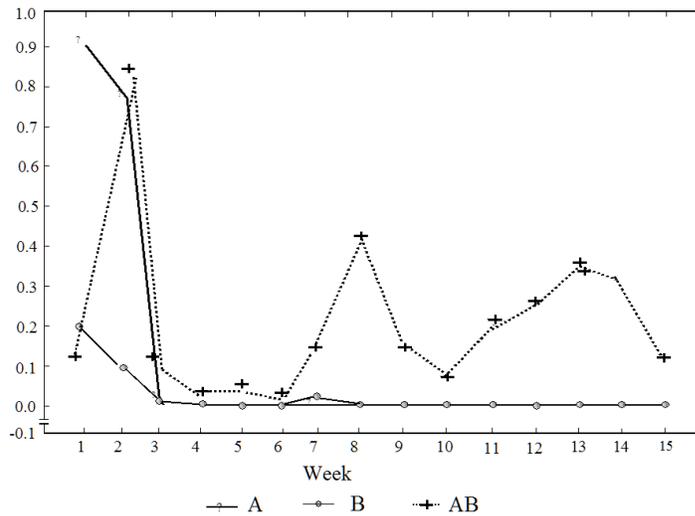


Figure 2: Plots showing the change in p -values for the significance tests associated with A Factor, B Factor and AB interaction in an univariate two-way ANOVA for each week.

| | B_1 | B_2 | B_3 | B_4 | B_5 |
|-------|-------|--------|--------|--------|---------|
| A_1 | 2.238 | -0.483 | -0.162 | -0.164 | -0.0224 |
| A_2 | 0.117 | -0.540 | -0.241 | -0.298 | -0.211 |

Table 1: Fixed-effect estimates considering A , B and AB components for each mean μ_{ij} in the two-way layout. The estimate of the standard error was 0.0373.

5 Discussion

The proposed model can be used for modelling growth trajectories using random regression coefficients in a two-way crossed experimental design, and also for modelling the fixed-effect component as is commonly done in agricultural studies. The family of proposed models also includes L -degree polynomial regression models on time, which are commonly referred to as growth models. The proposal considers the equal spacing and balanced sequences of responses over time, but it does not require balance in the experimental design. We have also assumed simple multivariate normal distributions among measurements and the random regression coefficients. In many practical situations where the objective is to study height data there are also some constraints which ought to be included in the model. Finally, the extension of the model to a more general case follows readily (see, e.g., Goldstein, 1995; Hand and Crowder, 1990; Verbeke and Molenberghs, 1997).

Acknowledgments

We are very grateful with the editor and with two anonymous referees for comments and suggestions for improving our work. Also we are in debt with Guillermo Cruz for this valuable support in preparing the final version.

References

- [1] Bryk, A.S.; Raudenbush, S. W. (1992) *Hierarchical Linear Models: Applications and Data Analysis Methods*. Sage Publications, Thousand Oaks CA.
- [2] Chatfield, C. (1995) *Problem Solving; A Statistician's Guide*, Second Edition. Chapman and Hall, London.
- [3] Cox, D.R.; Snell, E.J. (1981) *Applied Statistics*. Chapman and Hall, London.
- [4] Crowder, M.J.; Hand, D.J. (1990) *Analysis of Repeated Measures*. Chapman and Hall, London.
- [5] Daniel, C.; Wood, F. S. (1980) *Fitting Equations to Data*. John Wiley, New York.
- [6] Dempster, A.P.; Laird, M.N.; Rubin, D.B. (1977) "Maximum likelihood from incomplete data via the EM algorithm", *Journal of The Royal Statistical Society, Series B* **39**: 1–38.
- [7] Dempster, A.P.; Rubin, D.B.; Tsutakawa, R.K. (1981) "Estimation in covariance components models", *Journal of the American Statistical Association* **76**: 341–353.
- [8] Diggle, P.J.; Liang, K.Y.; Zeger, S.L. (1994) *Analysis of Longitudinal Data*. Oxford University Press, New York.
- [9] Fussillier, L. (1996) *Un Modelo para Analizar Curvas de Crecimiento*. Tesina, Especialización en Métodos Estadísticos, Universidad Veracruzana, Veracruz, México.
- [10] Geisser, S. (1980) "Growth curve analysis", in: A.M. Kshirsagar (Ed.) *Handbook of Statistics*, Vol. 1, North-Holland, Amsterdam: 89–115.
- [11] Gibbons, R.D. (2000) "Mixed-Effects models for mental health services research", *Health Services & Outcomes Research Methodology* **1**(2): 91–129.
- [12] Gibbons, R.D.; Bhaumik, D.K. (2001) "Weighted random-effects regression models with application to interlaboratory calibration", *Technometrics* **43**(2): 192–198.
- [13] Goldstein, H. (1986) "Multilevel mixed linear model analysis using iterative generalized least squares", *Biometrika* **73**: 43–56.
- [14] Goldstein, H. (1987) *Multilevel Models in Educational and Social Research*. Grimm, London.

- [15] Goldstein, H. (1989a) “Models for multilevel response variables with an application to growth curves”, in: D.R. Bock (Ed.) *Multilevel Analysis of Educational Data*, Academic Press, New York.
- [16] Goldstein, H. (1989b) “Restricted unbiased iterative generalized least squares estimation”, *Biometrika* **76**: 622–623.
- [17] Goldstein, H. (1995) *Multilevel Statistical Models*, Second Edition. Halsted Press, New York.
- [18] Hand, D.J.; Crowder, M.J. (1996) *Practical Longitudinal Data Analysis*. Chapman and Hall, London.
- [19] Hartley, H.O.; Rao, J.N.K. (1967) “Maximum-likelihood estimation for the mixed analysis of variance model”, *Biometrika* **54**: 93–108.
- [20] Henderson, C.R. (1986) “Recent developments in variance and covariance estimation”, *Journal of Animal Sciences* **63**: 208–216.
- [21] Hocking, R.R. (1985) *The Analysis of Linear Models*. Brooks/Cole, Monterey, CA.
- [22] Hui, S.L. (1984) “Curve fitting for repeated measurements made at irregular time-points”, *Biometrics* **40**: 691–697.
- [23] Jenrich, R.I.; Schluchter, M.D. (1986) “Unbalanced repeated-measures models with structured covariance matrices”, *Biometrics* **42**: 805–820.
- [24] Kreft, I.G.G.; de Leeuw, J.; van der Leeden, R. (1994) “Review of five multilevel analysis programs: BMDP-5V, GENMOD, HLM, ML3, VARCL”, *The American Statistician* **48**: 324–335.
- [25] Langford, I.H.; Lewis, T. (1998) “Outliers in multilevel data”, *Journal of The Royal Statistical Society Series A* **161**: 121–160.
- [26] Laird, N.M.; Ware, J.H. (1982) “Random-effects models for longitudinal data”, *Biometrics* **38**: 963–974.
- [27] Lindsey, J. K.(1993) *Models for Repeated Measurements*. Oxford University Press, New York.
- [28] Longford, N. (1987) “A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed-models with nested random effects”, *Biometrika* **74**: 817–827.
- [29] Longford, N. T. (1993) *Random Coefficient Models*. Oxford University Press, New York.
- [30] Longford, N.T. (1995) “Random coefficient models”, in: G. Arminger, C.C. Clogg & M.E. Sobel (Eds.) *Handbook of Statistical Models for the Social and Behavioral Sciences*. Plenum Press, New York.

- [31] Mason, W.M.; Wong, G.Y.; Entwisle, B. (1983) “Contextual analysis through the multilevel linear model”, in: S. Leinhardt (Ed.) *Sociological Methodology*, Jossey-Bass, San Francisco.
- [32] McClean, R.A.; Sanders, W.L.; Stroup, W.W. (1991) “A unified approach to mixed linear models”, *The American Statistician* **45**: 54–64.
- [33] Ojeda, M. M.; Juárez C. S. (1996) “Biplot display for diagnostics in a two-level regression model for growth curves analysis”, *Computational Statistics and Data Analysis* **22**: 583–597.
- [34] Pan, H.Q.; Goldstein, H.; Guo, D. (1989) “A two-level cross-sectional model using grafted polynomials”, *Annals of Human Biology* **19**: 337–346.
- [35] Potthoff, P.F.; Rao, J.N.K. (1964) “A generalized multivariate analysis of variance model useful especially for growth curve problems”, *Biometrika* **51**: 313–326.
- [36] Prosser, R., Rassbash, J.; Goldstein, H. (1990) *ML3: Software for Tree Level Analysis*. Institute of Education, University of London, London.
- [37] Prosser, R., Rassbash, J.; Goldstein, H. (1991) *Data Analysis with ML3*. Institute of Education, University of London, London.
- [38] Ramsay, J.O.; Silverman, B.W. (1997) *The Analysis of Functional Data*. Springer-Verlag, New York.
- [39] Rao, C.R. (1965) “The Theory of least squares when the parameters are stochastic and its application to growth curves”, *Biometrics* **52**: 447–458.
- [40] Rawlings, J.O., Pantula, S.G.; Dicky, D.A. (1998) *Applied Regression Analysis: A Research Tool*. Second Edition. Springer-Verlag, New York.
- [41] Rencher, A. C. (2000). *Linear Models in Statistics*. John Wiley & Sons, New York.
- [42] Searle, S.R.; Casella, G.; McCulloch, C.E. (1992) *Variance Components*. John Wiley & Sons, New York.
- [43] Seber, G.A.F. (1984) *Multivariate Observations*. John Wiley & Sons, New York.
- [44] Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models and individual growth models. *Journal of Educational and Behavioral Statistics*. 24, No. 4, 323-355.
- [45] Shi, L.; Ojeda, M.M. (2004) “Local influence in multilevel regression for growth curves”, *Journal of Multivariate Analysis* **91**(2): 282–304.
- [46] Strenio, J.F.; Weisberg, H.I.; Bryk, A.S. (1983) “Empirical Bayes estimation of individual growth-curve parameters and their relationship to covariates”, *Biometrics* **39**: 71–86.

- [47] Verbeke, G.; Molenberghs (eds.) (1997) *Linear Mixed Models in Practice: A SAS-Oriented Approach*. Springer Verlag, New York.
- [48] Vonesh, E.F.; Chichilli, V.M. (1997) *Linear and Nonlinear Models for the Analysis of Repeat Measurements*. Marcel Dekker, New York.
- [49] Weissfeld, L. A.; Kshirsagar, A. M. (1992) “A modified growth curve model and its application to clinical studies”, *Australian Journal of Statistics* **34**: 161–168.
- [50] Yang, M., Goldstein, H.; Rasbash. J. (1996) *MLn Macros for Advanced Multilevel Modelling*. Institute of Education, University of London, London.
- [51] Ware, J. H. (1985) “Linear models for the analysis of longitudinal studies”, *The American Statistician* **39**: 95–101.