Multilevel Models in

Human Growth and Development Research

Huiqi Pan

Thesis submitted to the University of London

For the Degree of Ph.D

University of London

Institute of Education

Department of Mathematics, Statistics and Computing

March, 1995

Abstract

The analysis of change is an important issue in human growth and development. In longitudinal studies, growth patterns are often summarized by growth 'models' so that a small number of parameters, or the functions of them can be used to make group comparisons or to be related to other measurements. To analyse complete and balanced data, growth curves can be modelled using multivariate analysis of variance with an unstructured variance-covariance matrix; for incomplete and unbalanced data, models such as the two-stage model of Laird and Ware (1982) or the multilevel models of Goldstein (1987) are necessary.

The use of multilevel models for describing growth is recognized as an important technique. It is an efficient procedure for incorporating growth models, either linear or nonlinear, into a population study. Up to now there is little literature concerning growth models over wide age ranges using multilevel models.

The purpose of this study is to explore suitable multilevel models of growth over a wide age range. Extended splines are proposed, which extend conventional splines using the '+' function and by including logarithmic or negative power terms. The work has been focused on modelling human growth in length, particularly, height and head circumference as they are interesting and important measures of growth. The investigation of polynomials, conventional splines and extended splines on data from the Edinburgh Longitudinal Study shows that the extended splines are better than polynomials and conventional splines for this purpose. It also shows that extended splines are, in fact, piecewise fractional polynomials and describe data better than a single segment of a fractional polynomial.

The extended splines are useful, flexible, and easily incorporated in multilevel models for studying populations and for the estimation and comparison of parameters.

Acknowledgements

I wish to express my gratitude to my supervisor, Professor Harvey Goldstein, for his invaluable instruction, assistance and support through the course of this work.

I am greatly indebted to Dr Shirley G Ratcliffe, for being able to learn her expertise and to use the data of the Edinburgh Longitudinal Study on which she has been working for twenty-three years. I also thank her for her valuable comments and personal help.

I should like to thank Professor James M Tanner and Professor Michael Preece for their support and confidence in my ability to work on the Edinburgh Longitudinal data. I am also grateful to Professor Michael Healy for his valuable comments on the graft; to Professor Darrell Bock for providing the AUXAL program and the discussion on growth models; to Dr Catherine S Berkey for her experience and support.

Many thanks to the team of the Multilevel Models Project, to the members of the Department and the WHO Collaborating Centre on Child Growth and Development in London, for providing the facilities, support and encouragement.

I am grateful for the financial assistance which helped support this study, including Overseas Research Students Awards from university of London during the period 1991 - 1994.

The Shanghai Second University, Department of Biostatistics and Mathematics is gratefully acknowledged for allowing me leave of absence to carry out this research.

May also I thank my husband, Dr Ke Zheng Liang, for his care and support during the past years. I am forever indebted to my parents, my sisters for their practical help.

Contents

Tables

Chapter 4:

Chapter 5:

Figures

Chapter 4:

Chapter 5:

Chapter 1

Introduction

It is important to analyse longitudinal data in human growth and development studies (Goldstein, 1979; Plewis, 1985; Hauspie, Lindgren, Tanner and Spruch, 1991). This means that researchers in public health, psychology and social sciences need to collect data over time rather than to collect data at a particular point in time. The data collected in the former case are longitudinal and in the later case cross-sectional.

Longitudinal data hold hierarchical structure: the repeated measurements are clustered within each subject, furthermore subjects can be grouped further for example in schools. This clustering or grouping generally induces non-independence between population units so that statistical models based on assumptions of independence become invalid (Paterson and Goldstein 1991).

Multilevel modelling is a powerful statistical technique for analysing individuals as members of social groups and is especially useful for repeated measures data (Paterson and Goldstein, 1991). It is essentially a hierarchical linear model and an extension of ordinary multiple regression.

Bryk and Raudenbush (1987) presented a three-level hierarchical linear model for studying school effects on children's growth during the primary years. The level 1 units are the times or occasions, the level 2 units are the children and the level 3 the schools. This example considers two distinct features of the growth system: the structure of the mean or average growth trajectory and the nature of the deviations of the individuals growth trajectories from the population mean.

Goldstein (1987) described how multilevel models can be used for the efficient statistical modelling of longitudinal data with an example of children's height where six successive measures of height are made an 138 children between ages of six and eleven. The level 1 units are the measurement occasions within individuals and level 2 the individuals. Quadratic polynomials were fitted to the data with gender differences in intercept and growth rate. An average height growth curve was estimated together with the variance among individuals and variance among occasions. See Goldstein (1986a, 1989) for details.

It is worth noticing that using multilevel models we do not require the same number of measures for each individual nor that the measures are made at the same time or occasion while previous models did (Rao 1959; Elston and Grizzle, 1962). And also we should notice that polynomials are mostly used in hierarchical models to structure the mean trajectory or curve of growth and development but this is not necessary and that is the topic of this thesis.

1.1 Recent developments in the statistical handling of Longitudinal growth data

The longitudinal growth study will be introduced in two stages:

(1) Identifying those individual growth models which fit the number of measurements over a period of aging to present the underlying growth process.

(2) Structuring overall growth models for the population, that is, mean growth curves for a group or population and the variation in growth between individuals.

Generally, the models for overall growth have identical form to the models for individuals but with variation being further structured .

1.1.1 Individual Based Models

Healy (1989) outlined recent developments in the statistical handling of growth data. Many longitudinal studies are designed to investigate changes over time in a characteristic which is measured repeatedly for each study participant. A smooth curve is fitted to an individual's data and from this curve several features of growth, such as peak height velocity and onset of pubertal spurt can be studied (Tanner, Whitehouse, Marubini and Resele, 1976; Hauspie, Lindgren, Tanner and Spruch, 1991).

We consider first parametric models. The first class of parametric models are the linear models, such as polynomials, $y = a + bt + ct^2 + \ldots$; the Count curve, $y = a + bt + clogt$ (Count, 1943) and its extension $y = a + bt + clogt + d/t$ (Berkey and Reed, 1987) which are suitable for the first few years of life. Non linear models such as the Jenss curve, $y = a + bt - \exp(c + dt)$ (Jenss and Bayley, 1937) can be used for young children; the Gompertz curve, $y = a\exp[-\exp(b - ct)]$ and the logistic $y = a + b/(1 + \exp(c + dt))$ are able to model the adolescent spurt (Deming, 1957). These models are suitable only for limited age ranges and several authors have proposed combinations of curves. A double logistic model for combining pre-adolescent and adolescent growth was first published by Bock, Wainer, Peterson, Thissen, Murray, and Roche (1973) followed by a modification, the triple logistic model with nine parameters (Bock and Thissen, 1976). Recently the triple logistic model has been modified to the BTT model (Bock, Toit and Thissen, 1994). The Preece-Baines curve (Preece and Baines, 1978) uses only five parameters, and has been used extensively for fitting longitudinal data on height. The JPPS model of Jolicoeur, Pontier, Pernin and Sempe (1988) and the JPA2 model of Jolicoeur, Pontier and Abidi (1992), with eight parameters, have been shown to fit height curves for infants as satisfactorily as for older children. See Appendix A for the triple logistic model, BTT model, JPA2 model and the Preece-Baines curve.

These growth models are mainly for height measurement. A model for head circumference from birth to 18 years was proposed by Roche, Mukerjee and Guo (1986), that is, $y = a[1 - be^{-c(t-d)^3}]$. The three parameter model $y = a + b(t)^{0.5} + c\log t$ was used for early ages (Guo, Roche and Moore, 1988).

As pointed out by Goldstein (1979) and by Healy (1989), the weakness of any method which imposes a fixed algebraic form upon the fitted growth curve is that this form may be too rigid to model the true complexities of the growth process. This will be especially true of curves based on few parameters. Generally, models with three or four parameters are capable of describing only a part of the growth curve, like the Jenss model, the Count model and the Reed model for the period of infancy and childhood, or the logistic model for the adolescent period. Models with a large number of parameters (5-7) are more complex in their mathematical expression and can cope with a broader age interval. Increasing the number of parameters also increases the flexibility of the curve to describe more detailed features of the growth process. However, there is a price to pay for that: fitting a eight-parameter model like the JPA2 model of Jolicoeur, Pontier and Abidi (1992), for example, requires a large number of observations in order to obtain a reliable fit. Even with a large number of parameters, there may always be certain events or short-term variations in growth rate which will not be shown by the fitted curve, simply because the mathematical function does not allow for it, for example, small spurts in prepubertal growth (Hauspie, Lindgren, Tanner and Spruch, 1991).

Turning now to look briefly at non-parametric models. In attempting to impose less rigidity in fitting the curve two approaches have been considered in addition to the heuristic graphical procedure used by Tanner, Whitehouse and Takaishi (1966), that is: spline and kernel estimation. A q-spline function is a piecewise or segmented polynomial of degree q with q-1 continuous derivatives at the change points, that is, knots, with which great flexibility of shape can be obtained (Cox,1971; Silverman, 1985; Seber and Wild, 1989). Variable knot cubic splines have been used to fit height curves for children aged four to eleven years (Berkey, Reed and Valadian, 1983). The shape-invariant model by Stützle, Gasser, Molinari, Largo, Prader and Huber (1980) is something of an intermediate between the parametric approach and the nonparametric spline functions (Healy 1989). A kernel estimation procedure has been discussed by Gasser, Müller, Köhler, Molinari and Prader (1984), Gasser, Müller, and Mammitzsch (1985), which is capable of fitting not only the smoothed distance curve but also its velocity and acceleration curves and by which growth parameters can be specified clearly. The cubic smoothing splines by Largo, Gasser, Prader, Stützle and Huber (1978) give results very like those of kernel estimation. See Appendix B for the kernel estimation of Gasser, Müller, Köhler, Molinari and Prader (1984).

1.1.2 Population Based Models

All these above models are individual based, that is, based upon fitting separate curves to each individual. In fact, a subject can be regarded as a member of a population and therefore population-based approaches should be considered in order to estimate population quantities and to model directly the variations in growth parameters for individuals. This is an important technique for comparing growth in different populations, as well as for studying the effects of other factors such as environmental ones. In addition, we may use the information on the population to improve the estimates of each subject's own parameters.

Following the paper by Wishart (1938), multivariate approaches have been developed to study population curves using the technique of fitting a curve to each individual followed by a series of regression analyses using the estimated individual curve coefficients.

Early models by a number of authors were devoted to the estimation of polynomial growth curves and to the comparison of growth curves (Wishart, 1938; Leech and Healy, 1959; Rao, 1959; Elston and Grizzle, 1962; Potthoff and Roy, 1964).

Generally, the approach used by the authors above is to consider the vectors y_j of observations on the jth individual assumed to occur at a fixed set of ages and to have independent multivariate normal distributions with common mean vector $X\beta$ and covariance matrix V. Thus, with the usual notation,

$y_j \sim N(X\beta, V)$,

where X is the design matrix raised on the ages when individuals were measured and β a vector of coefficients to be estimated. However, this model is not suitable for the situation when individuals are measured at arbitrary or unique times. The difficulty is that usually we can not control the circumstances under which the measurements are taken, and there may be considerable variation among individuals in the number and timing of observations.

The Bayesian linear model of Fearn (1975)

This approach postulates a separate growth curve for each individual and that the n_j observations y_j on the jth individual are independently and normally distributed about the curve for that individual, given parameter vectors β_j and design matrices X_j. The first-stage of this model is to model

(1) $y_j | \beta_j, \sigma_j^2 \sim N(X_j\beta_j, \sigma_j^2 I)$.

The second-stage is then to model

(2) $\beta_j | \mu, C \sim N(\mu, C)$,

where the β_j's are supposed exchangeable and particularly that the β_j's are independently and normally distributed with common mean vector μ and covariance matrix C.

We combine (1) and (2) to give

$y_j | \mu, \sigma_j^2, C \sim N(X_j\mu, X_jCX_j^T + \sigma_j^2 I)$.

The general Bayesian linear model estimation procedure of Lindley and Smith (1972) is used. This can be handled by recent developed Markov Chain Monte Carlo methods (Gibbs Sampler methods) for general case. Gilks, Clayton, Spiegelhalter, Best, McNeil, Sharples and Kirby (1993) gives a review of applications of Gibbs sampling.

Recently, efficient maximun likelihood methods have been devised for the estimates of these growth models with an unbalanced data setting, in which the number and time intervals of measurement can vary from individual to individual. The two-stage model by Laird and Ware (1982) and the multilevel model developed by Goldstein (1986a, 1986b, 1987, 1989) allow a very general approach to the modelling of growth data. I shall not consider the Bayesian model further.

The linear model of Laird and Ware (1982)

If the jth individual has a $n_j \times 1$ vector, y_j, of responses, the growth curve model assumes that each individual also has a vector of growth curve parameters, b_j^*. Let X_j be a known $n_j \times r$ design matrix linking b_j^* to y_j. For measured, multivariate normal data, the two-stage growth curve model of Laird and Ware (1982) is:

Firstly, for each individual unit, j,

(1) $y_j = X_j b_j^* + e_j,$

where e_j is distributed as $N(0,R_j)$. Here R_j is an $n_j \times n_j$ positive-definite covariance matrix; it depends on j through its dimension n_j, but the set of unknown parameters in R_j will not depend upon j.

In the second stage, the b_j^* are assumed distributed as $N(\beta, D)$, independently of each other and of the e_j. Here D is a $r \times r$ positive-definite covariance matrix and β is the unknown population parameters.

Marginally, the y_j are independent normals with mean $X_j\beta$ and covariance matrix $R_j + X_j D X_j^T$ and this is the same general form as before with R_j replacing $\sigma^2 I$.

If $b_j = b_j^* - \beta$ centers the individual random effects at 0, the model can be expressed as before as

(2) $y_j = X_j\beta + X_jb_j + e_j$.

Laird and Ware proposed a unified approach to fitting it, based on a combination of empirical Bayes and maximum likelihood under Normality using EM algorithm (Dempster, Laird and Rubin, 1977). The advantages of this formulation is that the data need no longer be balanced.

Very similar is the linear model of Strenio, Weisberg and Bryk (1983) which includes covariates. Their approach consists of first deriving Bayesian estimates based on known variances following Lindley and Smith (1972) and then substituting maximum likelihood estimates for the unknown variances in the estimation formulas. All these authors assume the form $R_j = \sigma^2 I_{(n_j)}$.

An example of using the Laird and Ware model for incorporating the Reed linear model for height using a sample of 62 boys aged at 8 to 18 years was reported by Berkey, Laird, Valadian and Gardner (1989).

The Multilevel Models of Goldstein (1986b, 1987)

The multilevel models developed by Goldstein (1986b, 1987) incorporate the above models as special cases. In particular they can accommodate covariates which change over time and within-subject so that, for instance, the variance can change with age (Goldstein, 1986a, 1987).

The models have the same form of Laird and Ware. The random terms $X_ju_j + e_j$ are assumed multivariate normal and maximum likelihood (ML) estimates or restricted maximum likelihood (REML) estimates via iterative generalized least squares (IGLS) is applied.

The literature review of Multilevel Models is given in section 2.1. Examples of using Multilevel Models for fitting polynomials to a group of children aged from 6 to 11 years can be found in Goldstein (1986a) and to a group aged from 10 years to adult with multivariate responses in Goldstein (1989).

A related procedure is that of Bryk and Raudenbush (1987), which uses empirical Bayes estimates to give maximum likelihood estimates in the Normal case via EM algorithm. Longford (1987) provides another procedure using the Fisher scoring algorithm for maximum likelihood estimation. Details of comparisons of the various statistical packages are given by Kreft, de Leeuw and Kim (1990).

The Empirical Bayes approach of Berkey (1982a)

The Jenss model for the growth of a single child, j, with n_j observations is

$$y_j = \alpha_{0j} + \alpha_{1j}x_j - \exp(\alpha_{2j} + \alpha_{3j}x_j) + e_j,$$

where y_j is a $n_j \times 1$ vector of measures, x_j a $n_j \times 1$ vector of age and e is $n_j \times 1$ vector of random error. Assuming that the residuals are uncorrelated and normally distributed

$$e_j | \theta \sim N(0, \sigma_j^2 I),$$

where I is an identity matrix of dimension n_j and

$$\theta_j^T = (\alpha_{0j}\, \alpha_{1j}\, \alpha_{2j}\, \alpha_{3j}).$$

They assume that the distribution of θ on the population is multivariate normal $N(\mu, \Sigma)$. They further assume empirically that θ and σ^2 are independent.

The estimation is done as follows: firstly, the Jenss growth curve parameters for each subject are estimated by nonlinear least squares and thus provide population estimates of μ and Σ;

secondly, these estimates of μ and Σ are inserted into the full likelihood. The procedure is iterative and provides empirical Bayes estimates of the parameters and the fitted curves for each child.

The Empirical Bayes approach of Bock and Thissen (1980)

This is similar to the above model; the residuals and growth parameters are assumed to have multivariate normal distributions $N(0, \Sigma_e)$ and $N(\mu, \Sigma_\theta)$ respectively.

The example used is the triple-logistic; fitting it to longitudinal data of 66 boys and 70 girls from ages one to eighteen years. The focus is primarily the estimation and prediction of individual curves.

The nonlinear model of Berkey and Laird (1986)

These authors assume that the growth of an individual can be modelled by a nonlinear function $g(*, *)$, e.g. the Jenss model, that is,

$$y_j = g(\theta_j, x_j) + e_j,$$

where x_j is the $n_j \times 1$ vector of ages for the jth individual; θ_j is the $r \times 1$ vector of his or her growth parameters. The residual terms are identically independently distributed as $N(0, \sigma^2)$ and the parameter vectors are distributed as $N(\mu_j, D)$. Each component of the mean vector μ_j is assumed to be a linear function of a time-constant covariate, so that $\mu_j = Z_j\gamma$ where Z_j contains the appropriate covariate values.

The authors focus on estimating effects of covariates, such as sex and protein on the parameters of early childhood growth, while Berkey (1982a) and Bock and Thissen (1980) use this basic model without covariates ($\mu_j = \mu$). They consider three methods of estimating γ and D: the ordinary unweighted least square (OLS), the univariate (ML) and multivariate (MML) versions of weighted least squares; the MML method corresponds to maximum likelihood under the multivariate normal model. They found that the difference between the predicted curves from the three methods of estimation were consistently much smaller than those between sex and between protein group curves though OLS is clearly inferior to either ML or MML. The EM algorithm (Dempster, Laird and Rubin, 1977) is used for both ML and MML.

The AUXAL Program of Bock, Toit and Thissen (1994)

AUXAL is a program for analysis of longitudinal measurements of human stature. In addition to quantitative information about individual growth curves, the program provides estimates of average curves for growth. The BTT model, an extension of the triple logistic model, and the JPA2 model of Jolicoeur, Ponitier and Abidi (1992) are chosen for modelling stature over ages from 6 months to maturity. On the assumption that the population distribution is multivariate normal, Bayes modal estimation is used for fitting. The Fishing-scoring (Newton-Gauss) method is applied to the Bayes modal estimation in the program. The population distributions used in the Bayes modal procedure for genders are the estimates of the population means and covariance matrices on data from the Fels Longitudinal study, which were estimated by the maximum marginal likelihood method (Bock, 1989). The structure average is based on the estimates of individuals.

1.2 Problems in Modelling Mean Growth with Wide Age Ranges

The literature concerning the use of longitudinal growth models with populations based on wide age ranges can be found as follows:

(1) The model proposed by Berkey and Laird (1986), which incorporates on the Jenss curve for height with covariates of sex and protein, covering age 0.25 to 6 years.

(2) The multilevel models of Goldstein (1986a) using polynomials for height with a covariate of sex, covering age 6 to 11 years. The multivariate model of Goldstein (1989) for height covers ages from 10 to 18 years.

(3) The model presented by Berkey, Laird, Valadian and Gardner (1989), which is an application of the Laird and Ware (1982) model for Reed's extension model on height with a covariate (protein intake), covering age 8 to 18 years.

In population studies, to date there are no publications which dealt with a model which can handle longitudinal growth data with covariates involving a wide age range, e.g. from birth to adulthood. There is a need for a method of estimating the effects of covariates or of making comparisons between mean curves of two or more populations, especially for height, and head circumference, both of which are important measurements in human growth.

The lack of models for analysing growth curves across a wide age range can be ascribed to two problems. First is the fact that the large variation in growth between individuals, especially in adolescence, makes it difficult to find adequate growth functions. Secondly, there exist changes in the structure of the underlying function of the growth curve over a wide age range and the most flexible curves proposed, polynomials, cannot be used directly since they are globally determined by their values in any small interval and, hence, cannot model structure change even when they are required to be of high order. They will often fit badly at some ages, especially at the extremes of the range. An illustrative example of fitting for titanium heat data has shown that polynomials are inadequate for modelling such phenomena (Eubank, 1984; Seber and Wild, 1989). An example of the difficulty in fitting curves over a whole age range is at the pubertal stage for height measurement and early childhood for head circumference where there are rapid changes. If it is a non-linear curve it will typically require a large number of parameters which makes estimation difficult, or else may introduce fixed relationships between growth events which are unrealistic (Goldstein, 1979). Berkey (1982a) found that the problems concerning convergence and uniqueness of solutions in nonlinear models were still relevant in the model she used.

In addition, the question may be whether individual based models can be incorporated into a population based model properly. For example, the population based model of Berkey and Laird (1986) requires the multivariate normality of the population of parameter vectors obtained from the individual based model. Longitudinal growth studies usually have missing data and irregularity in the ages of measurement. Only recently have the efficient models of Laird and Ware (1982) and Goldstein (1986b, 1987) been devised for the analysis of growth data in the unbalanced data setting.

The multilevel model (Goldstein, 1986b, 1987) has important advantages: it can accommodate covariates which change over time, and within-subject residual terms which are more complex than those assumed by Berkey, Laird, Valadian and Gardner (1989). While this methodology can be extended to nonlinear models, linear models are easier to interpret, flexible and computationally relatively straightforward (Healy, 1989). Thus linear models, whether for describing an individual or a population, are easier to implement both theoretically and computationally than nonlinear models. In addition, a final difficulty with existing longitudinal growth models, such as the Preece-Baines model and the JPA2 model is that they have been developed for analysis of growth in height and are generally not applicable to other measurements. For example they may be quite inappropriate for fitting the dimension of head circumference, which shows little or no adolescent growth spurt (Hauspie, Lindgren, Tanner and Spruch, 1991).

Up to now, little has been achieved in seeking an appropriate linear growth model suitable for describing the growth pattern through early childhood to adulthood on the population. Potential candidates are the general Reed model and splines.

The General Reed Model

Berkey, Laird, Valadian and Gardner (1989) proposed that if measurements earlier than age 8 years are included in the analysis, the general Reed model should be used, namely

$$y_{ij} = b_{0j} + b_{1j}x_{ij} + b_{2j}\ln(x_{ij}) + b_{3j}/x_{ij} + b_{4j}/x_{ij}^2 + b_{5j}/x_{ij}^3 + b_{6j}/x_{ij}^4 + e_{ij}.$$

The variable y_{ij} is the *j*th subject's *i*th length measurement obtained at age x_{ij} years. Though this model has not been commonly pursued it remains linear and indicates that other terms could be considered when a simple polynomial is not adequate (see Goldstein 1986a).

Another approach is to fit separate curves to different sets of occasions with smooth joins where the curves meet. For example, the Jenss and Bayley curve might be joined to a logistic curve at about age 10 (see Goldstein 1979). This idea can be found in the ICP model of Karlberg (1989). But up to now no success has been achieved with smoothing joins for separate models due to numerical difficulties and the requirement for a large number of measurements in order to obtain good estimates.

The only linear model proposed to fit both early childhood and adolescence is the general eight-parameter Reed model by Reed and Berkey (1989). The model combines the five-parameter Reed model for early childhood,

$$y = c_1 + c_2x + c_3\ln(x) + c_4/x + c_5/x^2$$

and a separate five-parameter model for adolescence,

$$y = a_1 + a_3\ln(x) + a_4/x + a_5/x^2 + a_6/x^3,$$

where x is age and y is length or weight. The Reed model for childhood is splined to one for the adolescent period in such a way that the curve is continuous in distance and velocity at the age where the two models are joined. This eight-parameter Reed model has been fitted to two boys and two girls (Reed and Berkey, 1989).

Splines

In the real sense, splines are an evolution of classical parametric inference and bridge the gap between parametric and nonparametric models. Spline functions, or regression splines, can be expressed as a linear combination of basis functions that usually have a polynomial representation (Wold, 1974; Wegman and Wright, 1983). Thus spline functions can be computed based on a least squares approach.

Spline functions are generally defined to be piecewise polynomials of degree n whose function values and first $n - 1$ derivatives agree at the points where they join. The abscissae of these join points are called knots. Polynomials may be considered a special case of splines with no knots, and piecewise polynomials with fewer than the maximum number of continuity restrictions may also be considered spline functions. Spline functions will be referred to as splines in the text.

Splines possess the property of having local behaviour that is less dependent on their behaviour elsewhere (Cox, 1971). That is also a feature of experimental data or measurement data, such as growth data. Splines have been utilized for diverse purposes in Agriculture, Economics, Pharmacokinetics, Geophysics and Astrophysics because of their ability to provide simple approximate models for complicated phenomena which are either difficult or impossible to model precisely (Eubank, 1984). These successful applications give credibility to the statement (Poirier, 1973):

> "Spline functions, and, more generally, piecewise polynomial functions are the most successful approximating functions in use today, They combine ease of handling in a computer with great flexibility, and therefore particularly suited for the approximation of experimental data or design curve experiments".

Splines with fixed knots are straightforward to estimate using restricted least squares estimation, with continuity restrictions on joints (Buse and Lim, 1977), but deciding on the number and placing of knots and the degree of the polynomial pieces is still a problem. (Smith, 1979).

A simpler approach to these is the use of truncated polynomial, or '+' functions representation. The '+' function is defined as

$$\begin{aligned} u_+ &= u, \quad if \;\; u > 0 \\ &= 0, \quad if \;\; u \leq 0. \end{aligned}$$

Smith (1979) provided the framework for a unified statistical theory of spline regression with fixed knots and using the '+' function representation. In general, with $m - 1$ knots, $\xi_1 < \ldots < \xi_{m-1}$, and m polynomial pieces each of degree n, the spline is expressed as:

$$f(x) = \sum_{i=0}^{n} \phi_{i0} x^i + \sum_{j=1}^{m-1} \phi_{nj} (x - \xi_j)_+^n \; . \qquad (1.1)$$

In this case, f and its first $n - 1$ derivatives will be continuous. Splines with $n = 3$ are cubic spline and are commonly used. Wold (1974) pointed that splines with 2nd and 3nd degree are computationally simple and have sufficient flexibility for most purposes.

However, there are limited reports of successful growth curve fitting over wide age ranges. Variable knot cubic splines were explored by Berkey, Reed and Valadian (1983) for fitting height curves to the Boston Longitudinal Data with satisfactory results under age 11 years and were found not adequate for adolescence when systematic errors were found. Variable knot cubic splines lead to a nonlinear estimation and thus all the problems arising in a nonlinear regression are present.

It may be useful to fit polynomial pieces of different degrees. The construction of spline models with polynomial pieces of different degrees has been illustrated by an example of a cubic-quadratic-linear spline, that is, the first piece cubic, the second quadratic, and the third linear (Smith 1979) and three other examples (Gallant and Fuller, 1973), which will be discussed in section 2.

Little is known about using splines with '+' function for different degrees in growth studies. How can these be structured to be suitable for fitting growth curves over wide age ranges? Are they flexible enough to fit curves for different measurements of length? Could any other function be used together with polynomials to form splines? How can they be incorporated into multilevel models with covariates other than age for comparison of growth among groups?

This study will explore the possibility of using splines with the '+' function of different degrees to describe a wide variety of adolescent growth patterns; and will investigate the validity of the particular models and incorporate the models into a more general multilevel structure. Thus, this thesis is going to investigate growth models which are suitable for a wide age range - from infancy to adulthood, and which are applicable to measurements such as height and head circumference; in addition the model will seek to handle unbalanced longitudinal data with several covariates.

Chapter 2 presents a review of the two-level model and splines with '+' functions; Chapter 3 describes the new model and Chapter 4 presents the results on real data. The final chapter provides a summary and discussion.

Chapter 2

Theory and Literature Review

This chapter outlines, in 2 sections, the foundations for the study. The first section 2.1 is a discussion of the fitting of two-level models. Piecewise polynomials are described in section 2.2.

2.1 Multilevel Models

The review will be focused on the structure of a basic two-level model with a simple level 1 covariance structure first and then the general two-level model with a complex level 1 covariance structure. The models for longitudinal growth curve fitting are described, followed by examples of a two-level model.

In addition the estimation of parameters and standard errors will be addressed. More details of model structure can be found in the work of Goldstein (1986b, 1987, 1992) and Prosser, Rasbash and Goldstein (1991).

2.1.1 A Basic Two-level Model

A two-level model with a simple level 1 covariance structure is now described. The term multilevel refers to a hierarchical relationship among units: in a longitudinal growth study measurements of an individual are regarded as level 1 units and individuals are level 2 units. In a survey, for example, in an education system, students (level 1) are members of classes which are level 2 units. In the following context level 2 units are also termed groups. Suppose in J groups, Y_j is a $n_j \times 1$ vector of the response variable values for the unit j, X_j is an $n_j \times r$ matrix of the values of members in group j on a set of r explanatory variables. Typically the intercept X_{0ij} is in this set. A within-unit model for this group is given as

$$Y_j = X_j\beta_j + e_j,$$

where β_j is a $r \times 1$ vector of the coefficients for the jth group, and e_j is a $n_j \times 1$ vector of level 1 random terms. The between-unit model for the coefficients can be written as

$$\beta_j = Z_j\Gamma + u_j,$$

where Z_j is an $r \times q$ between-unit design matrix, Γ is a $q \times 1$ vector of the fixed coefficients, and u_j is an $r \times 1$ vector of level 2 random terms. Combining the within- and between-unit models for the jth group gives

$$Y_j = X_jZ_j\Gamma + X_ju_j + e_j. \tag{2.1}$$

The $X_jZ_j\Gamma$ term is called the fixed part and the latter two terms form the level 2 random part and level 1 random part respectively.

On the assumption of independence of level 1 and level 2 random terms and the e_{ij} in group j are independently distributed with an expected value 0 and a variance of σ_e^2, the variance of Y_j conditional on the fixed part can be expressed as

$$\Sigma_j = Var(X_ju_j + e_j) = X_j\Omega_{(2)}X_j^T + \sigma_e^2 I_{n_j},$$

where $\Omega_{(2)} = Var(u_j)$. Typically, multivariate normality is assumed for random terms. The assumption is usually made that level 2 random terms in u_j have a joint distribution with mean 0 and covariance matrix $\Omega_{(2)}$.

The model for the total J groups gives

$$\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_J \end{pmatrix} = \begin{pmatrix} X_1 & 0 & \dots & 0 \\ 0 & X_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & X_J \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_J \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_q \end{pmatrix} + \begin{pmatrix} X_1 & 0 & \dots & 0 \\ 0 & X_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & X_J \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_J \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_J \end{pmatrix},$$

namely,

$$Y = XZ\Gamma + Xu + e \tag{2.2}$$

The covariance matrix of the random part $Xu + e$ can be expressed as follows:

$$\Sigma = X \begin{pmatrix} \Omega_{(2)} & 0 & \dots & 0 \\ 0 & \Omega_{(2)} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \Omega_{(2)} \end{pmatrix} X^T + \begin{pmatrix} \sigma_e^2 I_{n_1} & 0 & \dots & 0 \\ 0 & \sigma_e^2 I_{n_2} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \sigma_e^2 I_{n_J} \end{pmatrix}.$$

Note that Σ is an $N \times N$ block diagonal matrix where $N = \sum_{j=1}^{J} n_j$.

This basic two-level model imposes restrictions in several ways: all explanatory variables that have appeared in the random part of the model have been included in the fixed part; and, only the intercept can be considered to be random at level 1. This is not always desirable and will be dealt with in the next section.

2.1.2 A General Two-level Model

The general two-level model now to be described is more flexible. For instance, one can fit a complex model to the fixed part and very simple model to the random part and perhaps, fit a complex level 1 covariance structure. A shift in notation will be introduced to express the general model.

Suppose there are r_2 explanatory variables random at level 2 and r_1 at level 1 and the corresponding matrices of these explanatory variables are denoted as $X_{(2)}$ and $X_{(1)}$ respectively. The sets of variables in $X_{(2)}$ and $X_{(1)}$ will often intersect. Both will often be submatrices of X and will usually contain the constant term which defines the intercept in the fixed part.

For group j, the general two-level model can be written as:

$$Y_j = X_j Z_j \Gamma + X_{(2)j} u_j + \begin{pmatrix} X_{(1)1j}^T & 0 & \cdots & 0 \\ 0 & X_{(1)2j}^T & \cdots & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & 0 & \cdots & X_{(1)n_j j}^T \end{pmatrix} \begin{pmatrix} e_{1j} \\ e_{2j} \\ \cdot \\ \cdot \\ \cdot \\ e_{n_j j} \end{pmatrix} = X_j Z_j \Gamma + X_{(2)j} u_j + X_{(1)j} e_j, \qquad (2.3)$$

where, u_j is a $r_2 \times 1$ vector and e_j is a $n_j r_1 \times 1$ vector. $X_{(2)j}$ is a $n_j \times r_2$ matrix and $X_{(1)j}$ is $n_j \times n_j r_1$.

Thus the covariance matrix of Y_j is

$$\Sigma_j = X_{(2)j} \Omega_{(2)} X_{(2)j}^T + X_{(1)j} (I_{n_j} \otimes \Omega_{(1)}) X_{(1)j}^T,$$

where $\otimes$ is the Kronecker product; $\Omega_{(2)}$ is the covariance matrix of the u_j and $\Omega_{(1)}$ of the e_{ij}. For J groups, Σ is a block diagonal matrix composed of the Σ_js.

MAXIMUM LIKELIHOOD ESTIMATION VIA IGLS

The iterative generalized least squares (IGLS) algorithm (Goldstein, 1986b, 1987) is applied to fitting the model. Goldstein (1986b) has shown that when the disturbances follow a multivariate normal distribution, IGLS estimates are maximum likelihood. Under Normality assumptions for the random terms, the loglikelihood function for Γ and Σ given Y is:

$$l(\Gamma, \Sigma | Y) = \ln|\Sigma| + (Y - XZ\Gamma)^T \Sigma^{-1} (Y - XZ\Gamma),$$

which is minimized using the IGLS algorithm (see Appendix C).

2.1.3 An Example for Growth Curves

In a longitudinal growth study individuals are measured repeatedly. In terms of two-level models, the individual is regarded as a level 2 unit and the measurements on an individual are level 1 units. A comprehensive exposition on the modelling of longitudinal data can be found in Goldstein (1986a, 1987, 1989) and fitting polynomials is suggested as a promising approach to summarize growth.

If y_{ij} denotes the ith measurement of person j at occasion i and t_{ij}, the within-unit equation of a two-level polynomial growth model, in this case a quadratic, is expressed as follows:

$$y_{ij} = \beta_{0j} + \beta_{1j}t_{ij} + \beta_{2j}t_{ij}^2 + e_{ij}, \qquad (2.4)$$

where e_{ij}'s are the level 1 random terms . It is assumed that the level 1 random terms for an individual are distributed independently with mean 0, $Var(e_{ij}) = \sigma_e^2$ and $Cov(e_{ij}, e_{i'j}) = 0$.

In order to obtain each person's unique curve, a between-person, that is, between-unit model is needed. A simple example is

$$\beta_{0j} = \gamma_1 + u_{0j},$$

$$\beta_{1j} = \gamma_2 + u_{1j},$$

$$\beta_{2j} = \gamma_3 + u_{2j},$$

where the γ's are the fixed parameters of the mean growth curve; the u's are level 2 random terms, departures from the overall means by each person's own parameters. We have

$$Var(u_{kj}) = \sigma_{uk}^2,$$

$$Cov(u_{kj}, u_{k'j}) = \sigma_{ukk'}^2.$$

Some characteristic, for example sex (denoted by Z_{sj}), which remains fixed across occasions, can be incorporated into the between-unit models to account for coefficient variability:

$$\beta_{0j} = \gamma_1 + \gamma_2 Z_{sj} + u_{0j},$$

$$\beta_{1j} = \gamma_3 + \gamma_4 Z_{sj} + u_{1j},$$

$$\beta_{2j} = \gamma_5 + \gamma_6 Z_{sj} + u_{2j}. \tag{2.5}$$

Combining the within-unit model in equation (2.4) with the between-unit model in equation (2.5) gives the model with a simple level 1 covariance structure as follows:

$$y_{ij} = (1, t_{ij}, t_{ij}^2)\begin{pmatrix}\beta_{0j}\\ \beta_{1j}\\ \beta_{2j}\end{pmatrix} = (1, t_{ij}, t_{ij}^2)\begin{pmatrix}1 & Z_{sj} & 0 & 0 & 0 & 0\\ 0 & 0 & 1 & Z_{sj} & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & Z_{sj}\end{pmatrix}\begin{pmatrix}\gamma_1\\ \gamma_2\\ \gamma_3\\ \gamma_4\\ \gamma_5\\ \gamma_6\end{pmatrix} + (1, t_{ij}, t_{ij}^2)\begin{pmatrix}u_{0j}\\ u_{1j}\\ u_{2j}\end{pmatrix} + e_{ij}. \tag{2.6}$$

The variance of the level 2 random part for y_{ij} is

$$\sigma_{u0}^2 + t_{ij}^2\sigma_{u1}^2 + t_{ij}^4\sigma_{u2}^2 + 2t_{ij}\sigma_{u01} + 2t_{ij}^2\sigma_{u02} + 2t_{ij}^3\sigma_{u12}$$

and the variance of level 1 random part is

$$\sigma_{e0}^2.$$

In the model of equation (2.6), γ_1, γ_2, γ_3, γ_4,γ_5 and γ_6 are the fixed parameters to be estimated; σ_{u0}^2, σ_{u1}^2, σ_{u2}^2, σ_{u01}, σ_{u02} are the random parameters at level 2 to be estimated and σ_{e0}^2 at level 1.

A model with a complex level 1 covariance structure is practically useful. The work of Goldstein (1987) has introduced such elaboration of modelling level 1 dispersion as a function of other characteristics, such as age, gender etc.

Suppose level 1 random terms are a linear function of time, the model in equation (2.6) can be further structured as:

$$y_{ij} = (1, t_{ij}, t_{ij}^2)\begin{pmatrix} 1 & Z_{sj} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & Z_{sj} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & Z_{sj} \end{pmatrix}\begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \\ \gamma_5 \\ \gamma_6 \end{pmatrix} + (1, t_{ij}, t_{ij}^2)\begin{pmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{pmatrix} + (1, t_{ij})\begin{pmatrix} e_{0ij} \\ e_{tij} \end{pmatrix}. \qquad (2.7)$$

Thus level 1 variation, a quadratic function of time, is given by:

$$\sigma_{e0}^2 + 2t_{ij}\sigma_{e0t} + t_{ij}^2\sigma_{et}^2.$$

Similarly, if level 1 random terms in equation (2.6) are assumed to be a linear function of Z where Z_s is a (0,1) dummy variable for sex. We have the level 1 component

$$(1, Z_{sj})\begin{pmatrix} e_{0ij} \\ e_{sij} \end{pmatrix},$$

and the level 1 variation is

$$\sigma_{e0}^2 + 2Z_{sj}\sigma_{e0s} + Z_{sj}^2\sigma_{es}^2.$$

Setting $\sigma_{es}^2 = 0$ so that the variation of level 1 is a linear function of sex:

$$\sigma_{e0}^2 + 2Z_{sj}\sigma_{e0s},$$

i.e. for $Z_{sj} = 1$ the variance is $\sigma_{e0}^2 + 2\sigma_{e0s}$ and for $Z_{sj} = 0$ is σ_{e0}^2 (Goldstein, 1987).

Using general notation the model in equation (2.7) can be expressed as

$$y_{ij} = (X_{0ij}, X_{1ij}, X_{2ij})\begin{pmatrix} 1 & Z_{sj} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & Z_{sj} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & Z_{sj} \end{pmatrix}\begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \\ \gamma_5 \\ \gamma_6 \end{pmatrix} + (X_{0ij}, X_{1ij}, X_{2ij})\begin{pmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{pmatrix} + (X_{0ij}, X_{1ij})\begin{pmatrix} e_{0ij} \\ e_{tij} \end{pmatrix},$$

that is,

$$y_{ij} = (X_{ij}Z_{ij})^T\Gamma + X^T_{(2)ij}u_j + X^T_{(1)ij}e_{ij}. \tag{2.8}$$

Fitting a polynomial of degree p is quite straightforward.

2.2 Splines

In this section multiphase, that is, change of phase models, or piecewise regression is to be described first (Cox,1971; Seber and Wild, 1989) followed by an example of a growth model, which was the first linear model proposed for growth in stature from birth to maturity (Reed and Berkey, 1989). Then attention will focus on reviewing the contribution to the literature of Smith (1979), Wold (1974), Eubank (1984), Gallant and Fuller (1973), Cox (1971) and Seber and Wild (1989) on piecewise polynomials.

2.2.1 Piecewise Regression

There exist situations where the trajectory to be studied is composed of several regressions formed by piecing together different curves over different intervals.

Suppose $a = \xi_0 < \xi_1 < \ldots < \xi_{m-1} < \xi_m = b$ and a regression relationship between y and x is

$$\begin{aligned} f(x;\beta;\xi) &= f_1(x;\beta_1), \quad && if \ \ \xi_0 < x \le \xi_1; \\ &= f_2(x;\beta_2), && if \ \ \xi_1 < x \le \xi_2; \\ &\ldots \\ &= f_m(x;\beta_m), && if \ \ \xi_{m-1} < x < \xi_m, \end{aligned} \tag{2.9}$$

where $f_j(x;\beta_j)$ $(j = 1,2,\ldots,m)$ are linear function of x. Here $E[y/x] = f(x;\beta)$ is called piecewise regression or multiphase regression; the m submodels are referred to as phase models or regimes; the ξ's as changepoints or joinpoints. Typically the end points of the intervals are unknown and must be estimated which leads to nonlinear estimation.

Models of the form (2.9) are intended to fit curves in which each submodel can be described by a simple parametric function such as a straight line and there may be fairly abrupt changes between regimes including noncontinuous and continuous cases.

In the continuous case, suppose for $j = 1,2,\ldots,m-1$, q_j is the first derivative of $f(x;\beta;\xi)$ with respect to x which is discontinuous at $x = \xi_j$, namely, $\partial^s f/\partial x^s$ is continuous at ξ_j for $s = 0,1,\ldots,q_j - 1$ but discontinuous for $s = q_j$. The following constraints are imposed to make the curves smooth at joinpoints:

$$f_j^{(s)}(\xi_j) = f_{j+1}^{(s)}(\xi_j), \quad j = 1,2,\ldots,m-1; \quad s = 0,1,\ldots,q_j - 1. \tag{2.10}$$

Thus, for fixed ξ, the model (2.9) with linear submodels becomes a linear regression model subject to linear constraints. There have been a variety of computational techniques for piecewise regression models. Maximum likelihood estimation is often used. A Bayesian framework can also be considered. Seber and Wild (1989) provided references.

The segment can be a straight line or other function, for example, Lerman (1980) fits the form $f(x) = x/(x-1)$ as one of submodels and also considers segments of the form $f(x) = \cos(x)$. Reed and Berkey (1989) have proposed a piecewise linear regression in which linear, logarithm and increasing powers of the reciprocal terms are included, which appears to be the most recent work using piecewise regression in the field of growth curve fitting.

THE LINEAR MODEL OF REED AND BERKEY (1989)

The linear model of Reed and Berkey (1989) is composed of the five-parameter model (Berkey and Reed, 1987) for early childhood,

$$f_1(t) = \phi_{01} + \phi_{11}t + \phi_{21}\ln(t) + \phi_{31}/t + \phi_{41}/t^2, \qquad t < \xi$$

and an extension of this five parameter model for adolescence,

$$f_2(t) = \phi_{02} + \phi_{12}\ln(t) + \phi_{22}/t + \phi_{32}/t^2 + \phi_{42}/t^3. \qquad t \geq \xi$$

where t is age (in month) and $f(t)$ is measurement of height. ξ is the joinpoint or the boundary between the two periods of childhood and adolescence.

If age is set to be near zero at conception and 1 at the joinpoint by the following linear transform,

$x = (t+9)/(\xi+9)$,

then the above Reed models can be expressed in the form of the piecewise regression :

$$f_1(x) = \beta_{01} + \beta_{11}x + \beta_{21}\ln(x) + \beta_{31}/x + \beta_{41}/x^2, \qquad x < 1.0, \qquad (2.11)$$

$$\begin{aligned} f_2(x) = {} & \beta_{01} + \beta_{11}x + \beta_{21}\ln(x) + \beta_{31}/x + \beta_{41}/x^2 + \\ & \beta_{02} - \beta_{11}x + \beta_{12}\ln(x) + \beta_{22}/x + \beta_{32}/x^2 + \beta_{42}/x^3, \qquad x \geq 1.0. \qquad (2.12) \end{aligned}$$

The residuals from these models are assumed to be independent and distributed as $N(0, \sigma^2)$.

With an indicator z, the equation (2.11) and (2.12) can be combined into a single function

$$\begin{aligned} f(x) = {} & \beta_{01} + \beta_{11}x + \beta_{21}\ln(x) + \beta_{31}/x + \beta_{41}/x^2 + \\ & \beta_{02}z - \beta_{11}zx + \beta_{12}z\ln(x) + \beta_{22}z/x + \beta_{32}z/x^2 + \beta_{42}z/x^3, \qquad (2.13) \end{aligned}$$

where,

$$\begin{aligned} z &= 0, \quad \textit{if} \quad x < 1.0 \\ &= 1, \quad \textit{if} \quad x \geq 1.0. \end{aligned}$$

To smooth the curves, the following continuity constraints are imposed on the function and its first derivative:

$f_1^{(s)}(1) = f_2^{(s)}(1), \quad s = 0, 1$

where

$f_1^{(0)}(1) = \beta_{01} + \beta_{11} + \beta_{31} + \beta_{41};$

$f_2^{(0)}(1) = \beta_{01} + \beta_{02} + \beta_{31} + \beta_{22} + \beta_{41} + \beta_{32} + \beta_{42};$

$f_1^{(1)}(1) = \beta_{11} + \beta_{21} - \beta_{31} - 2\beta_{41};$

$f_2^{(1)}(1) = \beta_{21} - \beta_{31} - 2\beta_{41} + \beta_{12} - \beta_{22} - 2\beta_{32} - 3\beta_{42}.$

These two constraints lead to the following two expressions:

$\beta_{02} = \beta_{11} - \beta_{22} - \beta_{32} - \beta_{42}$

and

$\beta_{11} = \beta_{12} - \beta_{22} - 2\beta_{32} - 3\beta_{42}.$

Substituting these two expressions for β_{02} and β_{11} in the ten-parameter model (2.13) and rearranging terms results in the following eight-parameter Reed model:

$$\begin{aligned} y = {} & \beta_0 + \beta_1 \ln(x) + \beta_2/x + \beta_3/x^2 + \beta_4(z + x - zx + zln(x)) + \\ & \beta_5(-2z - x + zx + z/x) + \beta_6(-3z - 2x + 2zx + z/x^2) + \beta_7(-4z - 3x + 3zx + z/x^3). \quad (2.14) \end{aligned}$$

Obviously, the advantage of the linear function with the indicator z is that the model can be fitted by ordinary least-squares linear regression, which is easier than using a nonlinear procedure.

It seems that a reasonable choice of joinpoint $x = 1$ can be achieved by ad hoc solutions such as choosing the model with the smallest weighted residual variance with the joinpoint located at a specified age. It is necessary to experiment with a series of the ages in order to make an optimal choice for the boundary age. Ages (in years) of 6.25, 7.25, 8.25 and 9.25 were considered for girls and 7.25, 8.25, 9.25 and 10.25 for boys (Reed and Berkey, 1989).

Up to now the eight-parameter Reed model is the first linear model valuable for describing the human growth curve from birth to maturity. Good fitting was reported for two girls and two boys in distance curves but not in velocity curves. No population growth study has been made by using this eight-parameter model.

2.2.2 Piecewise Polynomials

A piecewise polynomial regression, or a segmented polynomial regression is a special continuous case of piecewise regression with continuity constraints of various orders in which the individual phase models are polynomials. The terminology of piecewise polynomials, segmented polynomials and grafted polynomials are often used interchangeably.

Let us return to the continuous case of piecewise regression (2.9). We assume

$$y_j = f(x_j;\beta,\xi) + e_j, \quad (j = 1,2,\ldots,m),$$

where $E[e_j] = 0, Var[e_j] = \sigma^2$, and e's are independent. Suppose the jth submodel in piecewise regression (2.9) is a polynomial of degree p_j, the piecewise polynomials can be expressed as

$$f_j(x) = \sum_{i=0}^{p_j} \beta_{ij}x^i, \quad \xi_{j-1} < x \le \xi_j, \quad j = 1,2,\ldots,m. \tag{2.15}$$

Again we assume q_j continuous derivatives at joinpoints ξ_j, $j = 1,\ldots,m-1$, i.e.

$$\left[\frac{d^s}{dx^s}\sum_{i=0}^{p_j}\beta_{ij}x^i\right]_{x=\xi_j} = \left[\frac{d^s}{dx^s}\sum_{i=0}^{p_{j+1}}\beta_{i,j+1}x^i\right]_{x=\xi_j}, \quad s = 0,1,\ldots,q_j - 1. \tag{2.16}$$

SPLINES USING '+' FUNCTION REPRESENTATION

There are various representations of piecewise polynomials. However in this thesis the focus is on the general form of piecewise polynomial, using '+' function representation, i.e.

$$\begin{aligned} u_+ &= u, \quad if \;\; u > 0 \\ &= 0, \quad if \;\; u \le 0. \end{aligned} \tag{2.17}$$

With this '+' function, the above model (2.15) can be written as a linear combination of terms $1, x, x^2, \ldots, x^p (p = \max_j p_j)$ and terms of the form $(x - \xi_j)_+^r$ $(r = 1, 2, \ldots, p)$ in such a way that all the continuity constraints are taken care of implicitly:

$$f(x) = \sum_i \phi_{i0} x^i + \sum_j \sum_r \phi_{rj} (x - \xi_j)_+^r \tag{2.18}$$

It means that if no values of r appearing is less than 2, then f has continuous derivatives with respect to either ξ_j or x of order $0, 1, \ldots, r - 1$. $(\xi_j - x)_+^r$ is another form of '+' function which can be used to structure the model (Seber and Wild, 1989).

A q-spline is a special case of the the model (2.17), in which every polynomial phase model has degree q and there are $q - 1$ continuous derivatives of x at the joints. A cubic spline is a q-spline where $q = 3$.

We now review three examples of Gallant and Fuller (1973), which illustrate how polynomial phases of different degree in models (2.15) with constraints in (2.16) can be expressed by a '+' function as in the model (2.17).

The quadratic-quadratic model ($m = 2, p_1 = p_2 = 2$): this model is composed of the first quadratic $f_1(x)$ and the second quadratic $f_2(x)$,

$$f_1(x) = \beta_{01} + \beta_{11} x + \beta_{21} x^2, \qquad a \le x < \xi_1;$$

$$f_2(x) = \beta_{02} + \beta_{12} x + \beta_{22} x^2, \qquad \xi_1 \le x < b.$$

The continuity constraints (2.16) on the function and its first derivatives are,

$$\beta_{01} + \beta_{11}\xi_1 + \beta_{21}\xi_1^2 = \beta_{02} + \beta_{12}\xi_1 + \beta_{22}\xi_1^2$$

and

$$\beta_{11} + 2\beta_{21}\xi_1 = \beta_{12} + 2\beta_{22}\xi_1. \tag{2.19}$$

With these constraints, the β_{01} and β_{11} are then be substituted from $f_1(x)$ by the expressions in terms of other parameters in the above equations of constraints to give

$$f_1(x) = \beta_{02} + \beta_{12}x + \beta_{22}x^2 + (\beta_{21} - \beta_{22})(\xi_1 - x)^2.$$

that is,

$$f_1(x) = f_2(x) + (\beta_{21} - \beta_{22})(\xi_1 - x)^2.$$

Then $f_1(x)$ when $\xi_1 - x > 0$ can be reparametrized as

$$f_1(x) = \phi_1 + \phi_2 x + \phi_3 x^2 + \phi_4(\xi_1 - x)_+^2$$

and $f_2(x)$ when $\xi - x \leq 0$ can be written as

$$f_2(x) = \phi_1 + \phi_2 x + \phi_3 x^2,$$

where $\phi_i = \beta_{02}, \phi_2 = \beta_{12}, \phi_3 = \beta_{22}$, and $\phi_4 = \beta_{21} - \beta_{22}$.

Now, with the '+' function (2.17), $f_1(x)$ and $f_2(x)$ can be expressed by $f(x)$,

$$f(x) = \phi_1 + \phi_2 x + \phi_3 x^2 + \phi_4(\xi_1 - x)_+^2. \qquad (2.20)$$

The initial polynomial term $\phi_1 + \phi_2 x + \phi_3 x^2$ thus expresses the second quadratic $f_2(x)$ when $\xi_1 - x < 0$ and the initial polynomial term together with the '+' term $(\xi_1 - x)_+^2$ expresses the first quadratic $f_1(x)$ when $\xi_1 - x \geq 0$.

To express a quadratic-quadratic model either $(\xi - x)_+^2$ or $(x - \xi)_+^2$ can be used. If $(\xi_1 - x)_+^2$ was replaced by $(x - \xi_1)_+^2$ in (2.20), the initial polynomial term is the first quadratic $f_1(x)$ and the the initial polynomial term together with the '+' term is the second quadratic $f_2(x)$.

$$f(x) = \sum_i \phi_{i0}x^i + \sum_j \sum_r \phi_{rj}(x - \xi_j)_+^r \tag{2.18}$$

The quadratic-linear model ($m = 2, p_1 = 2, p_2 = 1$): this model consists of the first quadratic $f_1(x)$ and the second linear $f_2(x)$

$f_1(x) = \beta_{01} + \beta_{11}x + \beta_{21}x^2, \qquad a \le x < \xi_1;$

$f_2(x) = \beta_{02} + \beta_{12}x, \qquad \xi_1 \le x < b.$

The continuity constraints on function and first derivative are:

$\beta_{01} + \beta_{11}\xi_1 + \beta_{21}\xi_1^2 = \beta_{02} + \beta_{12}\xi_1$

and

$\beta_{11} + 2\beta_{21}\xi_1 = \beta_{12}.$

With these constraints, the β_{01} and β_{11} are then be substituted from $f_1(x)$ by the expressions in terms of other parameters in the above equations of constraints to give

$f_1(x) = \beta_{02} + \beta_{12}x + \beta_{21}(\xi_1 - x)^2,$

that is,

$f_1(x) = f_2(x) + \beta_{21}(\xi_1 - x)^2.$

Then $f(x)$ can be reparameterized as

$$f(x) = \phi_1 + \phi_2 x + \phi_3(\xi_1 - x)_+^2, \tag{2.21}$$

where $\phi_i = \beta_{02}$, $\phi_2 = \beta_{12}$, $\phi_3 = \beta_{21}$. The initial polynomial term $\phi_1 + \phi_2 x$ is thus the second phase, the linear model and the initial polynomial term together with the '+' term $(\xi - x)_+^2$ is the first quadratic $f_1(x)$.

If $(\xi_1 - x)^2_+$ was replaced by $(x - \xi_1)^2_+$ in (2.21), the initial polynomial term is the first linear and the the initial polynomial term together with the '+' term is the second quadratic to express a linear-quadratic model instead of the quadratic-linear model. This means that $(x - \xi_1)^r_+$ is not able to express this model without imposing constraints on the coefficients of β's.

The quadratic-quadratic-linear model ($m = 3, p_1 = p_2 = 2, p_3 = 1$): this model is composed of the first quadratic $f_1(x)$, the second quadratic $f_2(x)$ and the third linear function $f_3(x)$ as follows:

$$f_1(x) = \beta_{01} + \beta_{11}x + \beta_{21}x^2, \qquad a \le x < \xi_1;$$

$$f_2(x) = \beta_{02} + \beta_{12}x + \beta_{22}x^2, \qquad \xi_1 \le x < \xi_2;$$

$$f_3(x) = \beta_{03} + \beta_{13}x, \qquad \xi_2 \le x < b.$$

Similarly, subject to continuous constraints on the function and the first derivative, the quadratic-quadratic-linear mosel is expressed as

$$f(x) = \phi_1 + \phi_2 x + \phi_3(\xi_1 - x)^2_+ + \phi_4(\xi_2 - x)^2_+, \tag{2.22}$$

where $\phi_1 = \beta_{03}$, $\phi_2 = \beta_{13}$, $\phi_3 = \beta_{22}$, and $\phi_4 = \beta_{21} - \beta_{22}$, while **the linear-quadratic-quadratic model** can be written as

$$f(x) = \phi_1 + \phi_2 x + \phi_3(x - \xi_1)^2_+ + \phi_4(x - \xi_2)^2_+.$$

Thus, the functions $1, x, x^2, \ldots, x^p$, $(\xi_j - x)^r_+$, ($p = \max_j p_j$, $j = 1, 2, \ldots, m-1$, $r = 2, 3, \ldots, p$) form a basis for piecewise polynomial satisfying that the functions and the first derivatives are continuous at joints. These examples show that to impose continuous first derivatives on the model we can delete $(\xi - x)_+$, and so on for other constraints (Gallant and Fuller, 1973).

We now write the piecewise polynomial in equation (2.18) in a more general form to give

$$f(x) = \sum_{i=0}^{n} \phi_{i0} x^i + \sum_{j=1}^{m-1} \sum_{r=k_j}^{p_j} \phi_{rj} (x - \xi_j)_+^r, \tag{2.23}$$

where $n \le p, 0 \le k_j \le p_j$. The continuity constraints can be imposed on the model by setting the value k_j in this way: if $q_j - 1$ continuous derivative are required at ξ_j let k_j be q_j so that the terms $(\xi_j - x)_+^r$ with $r \le q_j - 1$ are dropped from the model.

A n-spline is a special case of the general model in the equation (2.23) when $k_j = q_j = n$, $j = 1, 2, \ldots, m-1$, whose function values and first $n - 1$ derivatives agree at the points where they join. Another special case is the unconstrained or discontinuous piecewise polynomial when $k_j = 0$, $j = 1, 2, \ldots, m-1$. The former case is the smoothest one and the latter is the most discontinuous. The general form of piecewise polynomial itself sometimes is termed a weak spline or deficient spline, or even spline. In spline terminology, the joinpoints are termed knots.

It is not generally possible to use the basis without imposing constraints on the parameters ϕ unless the degrees of polynomial phase models are either nonincreasing $(p_1 \ge p_2 \ge \ldots \ge P_m)$ or nondecreasing $(p_1 \le p_2 \le \ldots \le p_m)$. The form $(\xi_j - x)_+^r$ or $(x - \xi_j)_+^r$ can used respectively without extra constraints (Seber and Wild, 1989). For instance, if $(x - \xi_j)_+^r$ is used instead of $(\xi_j - x)_+^r$ to express a model with nonincreasing degrees we need to impose constraints that some parameters are zero or certain pairs of parameters sum to zero so that the spline can model the specified polynomial phases and is continuous on its function or its derivatives. An example of a cubic-quadratic-linear spline using the $(x - \xi_j)_+^r$ form and continuity guaranteed on its function was given by Smith (1979):

$$\begin{aligned} f(x) = {} & \phi_{00} + \phi_{10} x + \phi_{20}\left[x^2 - (x - \xi_2)_+^2\right] + \phi_{30}\left[x^3 - (x - \xi_1)_+^3 - 3\xi_1 (x - \xi_2)_+^2\right] + \\ & \phi_{11}(x - \xi_1)_+ + \phi_{21}\left[(x - \xi_1)_+^2 - (x - \xi_2)_+^2\right] + \phi_{12}(x - \xi_2)_+, \end{aligned}$$

while the use of the form $(\xi - x)_+^r$ in this case makes the presentation simpler, giving

$$f(x) = \phi_1 + \phi_2 x + \phi_3(\xi_1 - x)_+^3 + \phi_4(\xi_2 - x)_+^2.$$

Without extra continuity constraints neither '+' function $(\xi - x)_+^r$ nor $(x - \xi)_+^r$ alone can be used to express a model which mixed by nonincreasing and nondecreasing submodels, such as linear-cubic-quadratic-linear. In chapter 3 and 4 we will discuss whether the '+' function has sufficient flexibility for growth studies with wide age ranges.

B-SPLINES

Apart from the piecewise polynomials in (2.15) and splines in (2.23), B-splines are also used (Wold, 1974; Eubank, 1984).

Suppose with $m - 1$ knots $a < \xi_1 < \xi_2 < \ldots < \xi_{m-2} < \xi_{m-1} < b$ the spline function of degree n is obeying continuity conditions for the function itself and its first $n - 1$ derivatives. Most commonly, n equals three, a cubic spline which can be expressed as follows:

$$f_j(x) = \phi_{0j} + \phi_{1j}x + \phi_{2j}x^2 + \phi_{3j}x^3, \qquad \xi_{j-1} \le x < \xi_j; \quad (\xi_0 = a;\ \xi_m = b).$$

In terms of B-splines, for cubic splines:

$$f(x) = \sum_{t=-1}^{m+1} \lambda_t B_t(x). \tag{2.24}$$

The $B_t(x)$'s are defined by means of divided differences

$$B_t(x) = \sum_{k=t-2}^{t+2} (x - \xi_k)_+^3 / \prod_{\substack{s=t-2 \\ s \ne k}}^{t+2} (\xi_k - \xi_s), \tag{2.25}$$

where additional knots are defined by

$$\begin{aligned} \xi_k &= \xi_1 - (1-k)(\xi_1 - a), && if\ \ k \le 0 \\ &= \xi_{m-1} + (k - m + 1)(b - \xi_{m-1}), && if\ \ k \ge m. \end{aligned} \tag{2.26}$$

The definition of (2.24), (2.25) and (2.6) make the B-splines have the property

$$B_t(x) \equiv 0 \qquad \text{when} \quad x > \xi_{t+2} \quad \text{or} \quad x < \xi_{t-2}. \qquad (2.27)$$

An advantage is that the number of the unknown parameters (λ_t) of B-splines is the same as the number of free parameters in the spline function. This leads to the fitting B-splines as a linear problem once the positions of knots are specified. For example the number of free coefficients of a cubic spline is equal to $m(n+1) - n(m-1) = m + n = m + 3$, which is the same number of (λ_t) in (2.24). Another advantage is that with the property of (2.27) B-splines have computational efficiency when a large number of knots are specified. This is due to the heptadiagonal structure of the moment matrix (X^TX) in least squares (Wold, 1974; Eubank, 1984).

The computation of the polynomial coefficients can be obtained from the B-splines coefficients in equation (2.24) by solving the following equations:

$$f'''(\xi_j) = 6\phi_{3j} = \sum \lambda_t B_t'''(\xi_j),$$

$$f''(\xi_j) = 2\phi_{2j} + 6\phi_{3j}\xi_j = \sum \lambda_t B_t''(\xi_j),$$

$$f'(\xi_j) = \phi_{1j} + 2\phi_{2j}\xi_j + 3\phi_{3j}\xi_j^2 = \sum \lambda_t B_t'(\xi_j),$$

$$f(\xi_j) = \phi_{0j} + \phi_{1j}\xi_j + \phi_{2j}\xi_j^2 + \phi_{3j}\xi_j^3 = \sum \lambda_t B_t(\xi_j), \quad j = 1, 2, \ldots, m. \qquad (2.28)$$

We can see from the equations that B-splines are not straightforward for statistical interpretation on these polynomial coefficients.

SMOOTHING SPLINES

Smoothing splines are an approach to the estimation of a smooth curve with a certain degree of smoothness. We can establish features of f such as its maximum value or prediction intervals for y, although the model parameters have no particular interpretation.

Suppose the data set (x_i, y_i), with $a \leq x_i < \ldots < x_n \leq b$, are generated by a regression model

$$y_i = f(x_i) + e_i \tag{2.29}$$

with $E[e_i] = 0$, $Var[e_i] = \sigma^2$, $i = 1,2,\ldots,n$. The smoothing involves the choice of the $\hat{f}$ as the function f with m derivatives which minimizes

$$S(f) = \sum_{i=1}^{n} [y_i - f(x_i)]^2 + \lambda \int_a^b \left(\frac{df^m(x)}{dx^m} \right)^2 dx. \tag{2.30}$$

The smoothing spline approach is a nonparametric regression problem. The first term is a least squares term. The second term is a 'roughness penalty' which is used to filter out very local variability due to random error so that the curve $\hat{f}$ does not have too many bends. The parameter $\lambda > 0$ adjusts the relative weighting given to the error sum of squares and the roughness penalty. Thus it controls the degree of smoothing. If λ is too small, the spline will overfit. If λ is too large, the smoothing term dominates and removes not only the noise but also the 'signal'. The method of cross-validation for choosing λ has been an option for choosing λ. A full illustration of the approach can be found in Silverman (1985), Wegman and Wright (1983).

2.2.3 Computation for '+' Function Splines

Among the above presentation of splines, the '+' function representation (2.23) is clearly a very useful one since it converts the spline (segmented polynomial) problem into an ordinary multiple regression one (Wegman and Wright, 1983). We choose to use '+' function splines for their relative ease of computation and interpretation.

Algorithms reviewed in this subsection are where the coefficients in the piecewise polynomials have joinpoints specified. Some hypotheses for splines with special interpretations will be discussed.

WITH FREE JOINPOINTS

Piecewise polynomials with joinpoints to be estimated from the data become nonlinear in parameters and estimation by least squares is difficult because it is very difficult to find joinpoints that give an absolute minimum for the residual sum of squares, although there exist a few strategies for the optimal selection of number and position of the joinpoints. Gallant and Fuller (1973) illustrated the procedure of using Modified Gauss-Newton fitting for piecewise polynomials with '+' function representation. Ertel and Fowlkes (1976) reviewed some algorithms for piecewise linear regression and proposed three algorithms in this direction. More detailed illustrations and references are given by Seber and Wild (1989).

WITH FIXED JOINPOINTS

There are several algorithms for different bases of piecewise polynomials with joinpoints or knots specified. Generally speaking, with the basis of the piecewise polynomial in equation (2.15) least square techniques are available for estimation (Cox, 1971; Poirier, 1973; Buse and Lim, 1977) and with the basis of '+' function representation ordinary least squares can be easily applied (Smith, 1979; Seber and Wild, 1989). The cubic spline is used for purposes of illustration.

a) Using Restricted Least Squares Method

The basis of model (2.15) is a linear model subject to linear constraints. Cox (1971) has provided a comprehensive account of both the piecewise and spline regression problems and suggested two general algorithms for the continuous case with fixed knots, in which Golub's method is used for solving a general linear least squares problem with linear equality constraints. Poirier (1973) proposed a least squares cubic spline regression method (CSR). Buse and Lim (1977) proved that the Poirier's CSR method is equivalent to the restricted least square estimation (RLS) and applied RLS to the same data set analyzed by CSR. Here the restricted least squares method used by Buse and Lim (1977) is demonstrated.

The cubic splines in equation (2.15) can be expressed as follows:

$$y = f_j(x) = \beta_{0j} + \beta_{1j}x + \beta_{2j}x^2 + \beta_{3j}x^3, \quad \xi_{j-1} \le x \le \xi_j, \quad j = 1, \ldots, m, \tag{2.31}$$

with the continuity constraints on the function, the first derivative and the second derivative as following:

$$f_j(\xi_j) = f_{j+1}(\xi_j), \quad j = 1, \ldots, m-1 \tag{2.32}$$

$$f_j'(\xi_j) = f_{j+1}'(\xi_j), \quad j = 1, \ldots, m-1 \tag{2.33}$$

$$f_j''(\xi_j) = f_{j+1}''(\xi_j), \quad j = 1, \ldots, m-1. \tag{2.34}$$

The cubic spline then is a linear model with $4m$ parameters and $3(m-1)$ linear restrictions. It can be written in matrix format:

$$Y = X\beta + e \tag{2.35}$$

and

$$C\beta = g, \tag{2.36}$$

where Y is $n \times 1$ vector,$n = \sum_{j=1}^{m} n_j$; $X = Diag[X_1, \ldots, X_j]$ is a $n \times 4m$ block diagonal matrix

where X_j is an $n_j \times 4$ matrix of observations; β is $4m \times 1$ vector of coefficients; C is an constraints matrix of $3(m-1) \times 4m$ and g is a $3(m-1) \times 1$ vector, usually $g = 0$. Minimizing the sum of squared residuals of (2.35) subject to the constraints of (2.36) gives the RLS estimator for β as

$$\beta_R = \beta + (X^TX)^{-1}C^{-1}[C(X^TX)^{-1}C^T]^{-1}(g - C\beta), \tag{2.37}$$

where $\beta = (X^T X)^{-1} X^T Y$, the ordinary least squares estimator. Two additional restrictions can be imposed on the end points and more details are shown in the work of Poirier (1973) and Buse and Lim (1977).

However the restricted least squares procedure is cumbersome and a total of $4m$ parameters must be estimated for a cubic spline compared to the free number of parameters $m + 3$. So for a problem with $m = 3$ we need to estimate 12 parameters instead of 6 free parameters.

b) Using Ordinary Least Squares

Now we turn from the piecewise polynomial basis to the '+' function basis. It has been mentioned that with the '+' function representation, the number of parameters to be estimated is the number of free coefficients as the continuity constraints have already been imposed implicitly. The model can simply be expressed as an multiple regression

$$Y = X\beta + e, \tag{2.38}$$

where β is a $k \times 1$ vector of coefficients to be estimated, Y is a $n \times 1$ vector of responses, n is the total number of observations; X is a $n \times k$ design martix and e is a $n \times 1$ vector of residuals. The ordinary least squares estimator of β in equation (2.37) is

$$\beta = (X^T X)^{-1} X^T. \tag{2.39}$$

However, the parameterization using the '+' function can lead to ill-conditioned design matrices, for example, when many knots are specified or when the knots are improperly placed. Thus, the choice of the number and positions of the knot is an important and difficult problem. Wold (1974) gives a full illustration on this problem reviewing his experience with spline functions.

Choice of knots: Wold (1974) proposed that the knots in a spline function should not be seen as ordinary free parameters, but their specification should rather be thought of as analogous to the choice of functional type. Hence, the knots should be chosen to correspond to the overall behaviour of the data. Here are the rules of thumb suggested by him for assisting judgement in cubic spline cases:

1 Have as few knots as possible.

2 Have not more than one extremum point (maximum or minimum) and one inflexion point per interval.

3 Have extremum points centred in the intervals.

4 Have inflexion points close to knots.

Wold (1974) recommended simulations to investigate the knot placing problem in the actual case by adopting these rules of thumb and transformation of data before fitting of splines if the data are not polynomial-like.

Testing: using the '+' function representation, testing whether a knot can be removed is easily accomplished by the hypothesis

H_0: $\xi_j = 0$

with t-statistics from the output of a multiple regression.

Smith (1979) has fully discussed using ordinary least squares to fit splines with '+' function representation and carried out tests of hypotheses by using standard multiple regression procedures. In the following section, the term regression splines will be used for the splines with '+' function representation.

Chapter 3

Two-level Growth Models for Length

This section presents models for length, focusing on head circumference (HC) and supine length (SL) or height (HT) which are the most difficult to fit over wide age ranges because there is a steep increment at early ages in HC and a large pubertal spurt increment in supine length and stature. First, for the within-individual model, extended splines using the '+' function representation are proposed in section 3.1. Then these extended splines, are incorporated in a 2-level Model with covariates, that is, the population based model in section 3.2. Section 3.3 deals with estimation and prediction. Hypothesis testing is discussed in section 3.4 (see Goldstein, 1987). Finally model checking using estimates of residuals is discussed in section 3.5.

3.1 Within-individual Model: Extended Splines

Though the general form of the conventional piecewise polynomials (2.23) is a useful tool there are some problems when it is applied to fitting of growth curves, especially with wide age ranges:

# it is not suitable for data where behaviour is non-polynomial;

# as the degree of the polynomials become smaller when equation (2.23) is differentiated the curves become smoother, that is, fluctuate less. This cannot deal with the fact that velocity or acceleration curves are observed to have more fluctuation than the distance growth curve itself (Gasser, Köhler, Müller, Kneip, Largo, Molinari and Prader, 1984).

As reviewed in section 2.1.1 the piecewise regression of Reed and Berkey (1989) used the log and reciprocal terms with constraints on the function and first derivative for the two segments of the regression. The computation will become more complicated if constraints on second derivatives are considered. So far no attempt has been made to add the log or reciprocal terms into the '+' function splines of equation (2.23).

The extended splines proposed in this section will be the first attempt at fitting growth models. The extended splines are an extension of piecewise polynomials of equation (2.23) with terms other than powers, such as terms of log or reciprocal if necessary. They can be written as

$$f(x) = \sum_{i=0}^{p_0} \phi_{i0} x^i + \sum_{j=1}^{m-1} \sum_{r=k_j}^{p_j} \phi_{rj} (x - \xi_j)_+^r + \gamma_1 G_1 \ln(ax + b) + \gamma_2 G_2 / x, \tag{3.1}$$

where G_1 and G_2 are constant, either 1 or 0; $m - 1$ is the number of joints and m is the number of segments in the spline; a and b are given constants and the values of knots are also given being $\xi_1 < \xi_2 < \dots < \xi_{m-1}$.

where

$$\begin{aligned} (x - \xi_l)_+^3 &= 0, & & \text{if } x \le \xi_l \\ &= (x - \xi_l)^3, & & \text{if } x > \xi_l. \end{aligned}$$

The log or reciprocal terms are introduced into the extended splines of model (3.1) in order to:

* describe data which are not purely polynomial-like;

* allow more fluctuations in the velocity and acceleration curves than in the distance curve itself.

It is obvious that the extended splines of model (3.1) inherit the advantages of conventional splines of model (2.23).

The location of ζ and ξ and the value of p_j, m can be determined by the overall behaviour of the data assisted by the adoption of the rules of thumb suggested by Wold (1974). We will allow the model (3.1) to have both terms $(x - \xi_j)_+^r$ and$(\zeta_j - x)_+^r$ if they do not overlap so that the model is neither restricted by nonincreasing or nondecreasing degrees of polynomials in each phase model.

We require a limitation on the number of coefficients β to be estimated and we suggest this is no more than eight as there are only about thirty to forty observations on each individual. Eight parameters have been considered for measurement of stature by several authors (Bock and Thissen, 1976; Reed and Berkey, 1989; Jolicoeur, Pontier and Abidi, 1992). Four to six parameters will be experimented with head circumference measurements.

Extended Splines for Head Circumference (HC)

Early in infancy there is a rapid increase in head circumference (HC) while after infancy HC increases slowly at least until 18 years (Roche, Mukherjee, Guo and Moore, 1987). This suggests a more complex pattern of head circumference growth curve in early childhood than that at other age ranges.

Let y_{ij} denote the ith head circumference measurement of the person j at time t_{ij} (in years). Model for head circumference proposed is as follows:

$$y_{ij} = \beta_{0ij} + \sum_{k=1}^{p} \beta_{kj} t_{ij}^{k} + \beta_{p+1,j} \ln(12t_{ij} + 1) + \sum_{k=1}^{m-1} \beta_{p+1+k,j} (\zeta_k - t_{ij})_+^3 + \sum_{k=1}^{n-1} \beta_{p+m+k,j} (t_{ij} - \xi_k)_+^3,$$

$$\beta_{0ij} = \beta_{0j} + e_{ij}, \tag{3.2}$$

where

$$\begin{aligned} (\zeta_k - t)_+^3 &= 0, & &\mathit{if} \quad t \geq \zeta_k \\ &= (\zeta_k - t)^3, & &\mathit{if} \quad t < \zeta_k \end{aligned}$$

and

$$
\begin{aligned}
(t-\xi_l)^3_+ &= 0, && if\ \ t \le \xi_l \\
&= (t-\xi_l)^3, && if\ \ t > \xi_l,
\end{aligned}
$$

where p is the degree of the polynomial phase. The values of knots are given so that $\zeta_1 < \zeta_2 < \ldots < \zeta_{m-1} < \xi_1 < \xi_2 < \ldots < \xi_{n-1}$.In this model, the total number of knots is equal to $m + n - 2$. The curve consists of$m + n - 1$ segments, each of which has continuous functions and continuous first and second derivatives.

For boys measured from birth to 15 or so, the initial knots are chosen at 2 and 10 years according to the nature of the growth and the model with six coefficients is written as:

$$y_{ij} = \beta_{0ij} + \beta_{1j}t_{ij} + \beta_{2j}t_{ij}^2 + \beta_{3j}\ln(12t_{ij}+1) + \beta_{4j}(2-t_{ij})^3_+ + \beta_{5j}(t_{ij}-10)^3_+.$$

$$\beta_{0ij} = \beta_{0j} + e_{ij}. \tag{3.3}$$

Extended Splines for Height (HT)

We consider the model not including the measures at birth. Let person j be measured on response y at age t (in year). For the jth person on the ith measurement of height, two models A and B will be explored.

MODEL HT-A:

$$y_{ij} = \beta_{0ij} + \sum_{k=1}^{p} \beta_{kj}t_{ij}^k + \beta_{p+1,j}\ln(t_{ij}) + \sum_{k=1}^{m-1} \beta_{p+1+k,j}(\zeta_k - t_{ij})^3_+,$$

$$\beta_{0ij} = \beta_{0j} + e_{ij}. \tag{3.4}$$

The values of knots are set at $\zeta_1 < \zeta_2 < \ldots < \zeta_{m-1}$. In this model, there are $m-1$ knots and m segments, each of which has continuous function and continuous first and second derivatives.

The intial values of knots are set to 9, 11, 13, 15, 17 years when girls were measured after birth up to 18.5 years. Thus the model with eight parameters can be written as:

$$y_{ij} = \beta_{0ij} + \beta_{1j}t_{ij} + \beta_{2j}\ln(t_{ij}) + \beta_{3j}(9 - t_{ij})^3_+ + \beta_{4j}(11 - t_{ij})^3_+ + \beta_{5j}(13 - t_{ij})^3_+ + \beta_{6j}(15 - t_{ij})^3_+ + \beta_{7j}(17 - t_{ij})^3_+,$$

$$\beta_{0ij} = \beta_{0j} + e_{ij}. \qquad (3.5)$$

For boys measured after birth up to 18.5 years the intials values of knots are set to 9, 11, 13, 15, 17.5 years.

Model HT-B:

$$y_{ij} = \beta_{0ij} + \beta_{1j}t_{ij} + \beta_{2j}\ln(t_{ij}) + \beta_{3j}1/t_{ij} + \sum_{k=1}^{n-1} \beta_{3+k,j}(\zeta_k - t)^3_+,$$

$$\beta_{0ij} = \beta_{0j} + e_{ij}. \qquad (3.6)$$

The values of knots are set at $\xi_1 < \xi_2 < \ldots < \xi_{n-1}$. There are $n - 1$ knots and n segments, each of which has continuous function and continuous first and second derivatives. The initial values of knots are 9, 11, 13 and 17 years when girls were measured after birth up to 18.5 years and the model is written as follows:

$$y_{ij} = \beta_{0ij} + \beta_{1j}t_{ij} + \beta_{2j}\ln(t_{ij}) + \beta_{3j}1/t_{ij} + \beta_{4j}(9 - t_{ij})^3_+ + \beta_{5j}(11 - t_{ij})^3_+ + \beta_{6j}(13 - t_{ij})^3_+ + \beta_{7j}(17 - t_{ij})^3_+,$$

$$\beta_{0ij} = \beta_{0j} + e_{ij}. \qquad (3.7)$$

The extended splines shown above are useful for many purpose on the grounds of simplicity and flexibility and possess the property of having a continuous function and continuous first and second derivatives without extra constraints. The form of these functions makes them suitable for use in multilevel models for descriptions of the trajectories of the growth and the variation between individuals together with covariates.

3.2 Between-individual Model

In terms of a 2-level model, individuals are regarded as level 2 units and measurements as level 1 units in longitudinal growth data. The functions (3.2), (3.4) and (3.6) are the within-individual model, in which the term e_{ij} is referred to the level 1 'residual' for the ith measurement in the jth individual. The between-individual model can be expressed as:

$$\beta_{0j} = \gamma_{00} + u_{0j},$$

$$\beta_{1j} = \gamma_{10} + u_{1j},$$

$$etc. \qquad (3.8)$$

where γ_{00} is the mean intercept and γ_{10} is the mean linear growth rate.

We assume that the level 1 random term e_{ij} is distributed independently from each of the level 2 random terms and we have

$$E(e_{ij}) = Cov(e_{ij}, e_{kj}) = E(u_j) = 0. \qquad (3.9)$$

Berkey (1982b) pointed out that the above assumptions may be questioned as there are correlations between errors when measurements are a very short time apart due to random fluctuation in growth rate. Berkey (1982b) and Goldstein, Healy and Rasbash (1994) presented the evidence that it is reasonable to assume uncorrelated errors when the observed ages are well spread out in time. El Lozy (1978), Preece and Baines (1978) and Jolicoeur, Pontier, Pernin and Sempe (1988) have made this assumption. In our study the target ages are similar to those of the above papers and we shall make this assumption also.

The variance of u_{0j} and u_{1j} and their covariance are denoted by σ^2_{u0}, σ^2_{u1} and σ_{u01} respectively.

The level 2 variances σ^2_{u0}, σ^2_{u1},... are examined to find out which coefficients should be considered random. The coefficients of the intercept β_{0j}, and slope β_{1j} are usually treated as random variables at level 2. We may wish to attempt to account for between-individual variation in terms of one or more features Z of individuals, for example gender, and the model can be extended by including such explanatory variables to give the following:

$$\beta_{0ij} = \gamma_{00} + \gamma_{01}Z_{1j} + u_{0j} + e_{ij},$$

$$etc. \qquad (3.10)$$

Likewise the slope, etc. coefficients can be modelled as functions of level 2 explanatory variables. Other features such as karyotype or protein intake may also be included in the between-individual model. Separate coefficients may be fitted for different subgroups, for example males and females, and this should be determined by knowledge of the subject matter.

The level 1 variances can be structured. Suppose we assume that the level 1 random term is a linear function of age t_{ij} and we specify the level 1 random component as

$$e_{ij} = e_{0ij} + e_{tij}t_{ij}, \qquad (3.11)$$

by setting the coefficient of t_{ij} random at level 1. Thus the level 1 variance is now a quadratic function of t_{ij}, namely

$$\sigma^2_{e0} + 2t_{ij}\sigma_{e0t} + t^2_{ij}\sigma^2_{et}, \qquad (3.12)$$

Also we can structure the level 1 random term as a function of gender (see Goldstein,1987).

Using general notation the model in (3.2) or (3.4) or (3.6) together with their level 2 and level 1 model (3.10) and (3.11) can be expressed as

$$y_{ij} = (X_{ij}Z_{ij})^T\Gamma + X^T_{(2)ij}u_j + X^T_{(1)ij}e_{ij}, \qquad (3.13)$$

where X_{ij} is a general notation for the known design variables such as 1, t_{ij}, ... and covariates; the $X^T_{(2)j}$ and $X^T_{(1)j}$ are rows containing the explanatory variables whose coefficients are random at level 2 and level 1 respectively.

3.3 Estimation and Predicted Values

Most statistical packages do not provide efficient estimates of the parameters in complex multilevel models. The Iterative Generalized Least Square (IGLS) algorithm (Goldstein, 1986b, 1987) is implemented in the ML3 software produced by the Multilevel Models Project, Institute of Education (Prosser, Rasbash and Goldstein, 1991). ML3-E version 2.3 was used for the computation in the study.

The computer package of HLM3 (Bryk, Raudenbush and Congdon, 1993) uses EM algorithm of Bryk and Raudenbush (1987) and the VARCL program produced by Longford (1987) uses the Fisher scoring algorithm. These can not fit models with a complex level 1 structure.

Mean Predicted Values

The estimation for fixed and random parameters using the IGLS algorithm can be found in Appendix C. The estimates of the fixed parameters, $\hat{\Gamma}$, are used to predict the mean response variable values for a given set of explanatory variables:

$$\hat{Y} = XZ\hat{\Gamma}. \tag{3.14}$$

Individual Predicted Values

The estimates of level 2 residuals u_j are given:

$$\hat{u}_j = (X_{(2)j}\hat{\Omega}_{(2)})^T \hat{\Sigma}_j^{-1} \tilde{Y}_j, \tag{3.15}$$

where $\tilde{Y}_j$ denotes the vector of the total residuals for the individual j, $\hat{\Omega}_{(2)}$ is the estimates of covariance of level 2 residuals and $\hat{\Sigma}_j$ is the estimates of the covariance of the Y_j. Thus predicted values for the individual j can be expressed as

$$X_jZ_j\hat{\Gamma} + X_{(2)j}\hat{u}_j. \tag{3.16}$$

3.4 Hypothesis Testing with Fixed Coefficients and Nested Models

Hypothesis Testing with Fixed Coefficients

We decided on the initial knots in the extended splines according to rules of thumb suggested by Wold (1974). We may change the location of the knots in order to find an appropriate model. A hypothesis concerning p elements of Γ can be formulated and tested using the estimated standard errors. A hypothesis about p elements of Γ is written:

H_0: $C\Gamma = K$

For example, the hypothesis with the first p elements may be:

$\gamma_1 = \gamma_2 = \dots = \gamma_p = 0$

and the H_0 is written:

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \dots & & & & \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ . \\ . \\ . \\ \gamma_p \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ . \\ . \\ . \\ 0 \end{pmatrix}.$$

ML3 will provide the quantity $(C\hat{\Gamma} - K)^T [C(XZ)^T \hat{\sum}^{-1} (XZ)C^T]^{-1} (C\hat{\Gamma} - K)$ and we refer it to the χ^2 distribution with p degrees of freedom.

Hypothesis Testing with Nested Models

The loglikelihood test statistic is a means of testing hypotheses about nested multilevel models. Suppose we have fitted model 1 and need to check whether one more knot is necessary, say model 2. Model 1 is nested in model 2, with loglikelihood functions l_1 and l_2

with number of parameters d_1 and d_2 correspondingly. The hypothesis of interest is: extra parameters are zero and the likelihood test statistic is

$$D = 2(l_1 - l_2)$$

where D has approximately a χ^2 distribution with $d_2 - d_1$ degrees of freedom in large samples. See Appendix 2.1 of Goldstein (1987) for details.

3.5 Checking Model Assumption

The estimates of residual terms can be used to check whether the underlying assumptions are adequate, e.g., to check for the gaussianity of the residual distribution. They can also be used to study whether further explanatory variables should be added to the model by plotting the corresponding residuals against those variables.

The variance of the residuals for a given level can be partitioned into two components, diagnostic variance and comparative variance. For the level 1 residual, the decomposition can be given as

$$Var(e_{ij}) = Var(\hat{e}_{ij}| e_{ij}) + Var(\hat{e}_{ij}). \qquad (3.17)$$

where $Var(\hat{e}_{ij}| e_{ij})$ is comparative variance and $Var(\hat{e}_{ij})$ diagnostic variance. The diagnostic variances can be used to standardize the estimated residuals for model-checking purposes. A further discussion of residuals is given in chapter 2 and 3 of Goldstein (1987).

Chapter 4

Examples

In this chapter the extended splines proposed in chapter 3 are applied to real longitudinal data sets of head circumference and height to model two-level structures with covariates. Estimates of random parameters and fixed parameters are given together with residual plots. Likelihood ratio tests and large sample 'Wald' tests (Pfeffermann and LaVange, 1989) based on contrasts (section 3.4) are used. Section 4.1 deals with head circumference and 4.2 height.

4.1 Modelling Head Circumference (HC)

Head circumference is an important part of the growth clinic examination. For example, in studying children with sex chromosomal aneuploidy, head circumference (HC) is an important factor in the anthropometric follow-up of these patients as a possible predictor of later cognitive ability (Ratcliffe, Masera, Pan and McKie, 1994).

The research on longitudinal data of head circumference was focused on obtaining reference data. Reference data for head circumference and 1-month increments from 1 to 12 months of age in the Fels Longitudinal Study were provided by Guo, Roche and Moore (1988). A three-parameter linear model was fitted to the head circumference data for each individual and the estimated head circumference data were used to calculate the reference. Roche, Mukherjee and Guo (1986) presented a four-parameter nonlinear model for head circumference from birth to 18 years. The four-parameter model was used to smooth the raw mean and the standard deviation for each age group in the Fels Longitudinal Study (Roche, Mukherjee, Guo and Moore, 1987). Recent work of Ratcliffe, Masera, Pan and Mckie (1994) studied karyotype effect on head circumference in the Edinburgh Longitudinal Study. Extended splines were used to fit data for each individual and the estimates were used to calculate mean and standard deviation for several age groups of controls and chromosome abnormal children. No other papers in the literature have dealt with population data in studying head circumference with abnormal karyotype as a covariate by using a random-effect model.

DATA

The subjects include 31 children with sex chromosome abnormalities (10 XYY, 11 XXY and 10 XXX) and 163 controls (83 XY and 60 XX), who were at least 14 years of age and had been identified by cytogenetic screening of consecutive liveborn infants between 1967 and 1979. Chromosomally normal male and female infants were recruited as controls between 1972 and 1976 from the two hospital in which the cytogenetic survey was being carried out (Ratcliffe and Paul 1986). Twins and low birth weight children were not included in this study. Details can be found in Ratcliffe, Masera, Pan and McKie (1994). Tables 4.1.1 to 4.1.3 show the number of measurements and cross-sectional means by age group, where HC measurements are in cm and age in years.

Table 4.1.1 The Number of measures (HC) by karyotype

Karyotype	Individuals	Total Measures	Mean Measures per individual
XY	83	2522	30.39
XX	60	1781	29.68
XYY	10	261	26.10
XXY	11	306	27.82
XXX	10	285	28.50

Table 4.1.2 Mean HC (cm) and number of measures by age group of controls

	XY			XX		
Age group	Mean	S.D.	N	Mean	S.D.	N
0.00	35.58	1.11	83	34.80	0.85	60
0.25+	41.90	1.36	85	40.58	0.99	58
0.50+	44.80	1.45	80	43.50	1.21	57
0.75+	46.65	1.32	73	45.46	1.14	52
1.00+	47.99	1.33	82	46.78	1.26	58
1.50+	49.51	1.36	80	48.03	1.23	58
2.00+	50.46	1.49	81	49.12	1.28	56
2.50+	51.14	1.40	73	49.88	1.33	52
3.00+	51.52	1.41	83	50.34	1.31	57
3.50+	51.97	1.45	82	50.77	1.26	59
4.00+	52.38	1.36	85	51.10	1.25	55
4.50+	52.70	1.40	76	51.35	1.22	64
5.00+	53.09	1.34	83	51.83	1.21	55
5.50+	53.09	1.52	78	51.88	1.14	54
6.00+	53.47	1.53	75	52.33	1.31	62
6.50+	53.45	1.36	79	52.27	1.14	52
7.00+	53.70	1.39	79	52.65	1.34	57
7.50+	53.95	1.46	78	52.80	1.26	55
8.00+	54.12	1.42	78	52.88	1.34	57
8.50+	54.27	1.44	76	53.27	1.18	52
9.00+	54.39	1.45	84	53.37	1.31	61
9.50+	54.58	1.44	76	53.48	1.21	54
10.00+	54.81	1.46	83	53.77	1.17	55
10.50+	54.79	1.42	78	53.75	1.38	55
11.00+	55.04	1.48	80	54.11	1.30	56
11.50+	55.09	1.48	78	54.29	1.40	58
12.00+	55.42	1.48	77	54.33	1.25	55
12.50+	55.35	1.40	79	54.47	1.31	49
13.00+	55.60	1.59	75	54.55	1.34	55
13.50+	55.91	1.44	67	54.76	1.26	56
14.00+	56.15	1.60	80	54.90	1.47	53
14.50+	56.46	1.51	80	55.05	1.46	51

Table 4.1.3 Mean HC (cm) and number of measures by age group of cases

		XYY			XXY			XXX	
Age group	Mean	S.D.	N	Mean	S.D.	N	Mean	S.D.	N
0.00	35.22	1.97	10	34.31	1.14	11	33.64	1.43	10
0.25+	43.24	2.87	11	40.81	1.15	6	40.47	1.93	8
0.50+	44.55	1.04	7	43.42	0.97	11	41.73	2.05	8
0.75+	46.96	2.18	5	45.49	1.20	10	43.87	1.80	7
1.00+	47.84	1.88	9	46.57	1.36	10	45.55	1.45	12
1.50+	49.14	1.87	11	48.06	1.28	12	46.64	1.27	9
2.00+	50.58	1.70	9	49.00	1.43	10	47.65	1.28	11
2.50+	51.46	2.07	6	49.65	1.51	10	48.42	1.38	10
3.00+	51.43	1.67	11	49.51	1.39	6	49.17	1.80	10
3.50+	52.86	1.44	8	50.50	1.25	11	48.94	1.07	9
4.00+	52.12	1.98	10	50.48	1.51	7	49.44	1.43	11
4.50+	53.08	1.53	7	51.00	1.20	9	49.71	1.84	6
5.00+	53.17	1.99	9	50.98	1.48	8	50.19	1.35	11
5.50+	52.80	1.91	9	51.27	0.94	8	49.71	1.16	7
6.00+	53.32	1.99	9	51.25	1.08	9	50.33	1.63	9
6.50+	54.01	1.84	9	52.23	1.38	11	51.13	1.34	8
7.00+	53.10	1.87	8	51.64	1.95	5	50.55	1.09	7
7.50+	54.23	1.77	8	52.15	0.94	9	51.36	1.38	11
8.00+	53.96	1.96	11	52.43	1.34	10	50.48	0.94	6
8.50+	54.43	1.82	6	52.88	1.04	9	50.95	1.17	7
9.00+	54.30	1.95	10	52.97	1.39	11	51.63	1.41	11
9.50+	54.45	2.00	6	53.00	0.76	8	51.48	1.79	7
10.00+	54.08	1.91	10	53.21	1.54	13	52.03	1.56	8
10.50+	54.95	1.69	4	53.37	1.45	10	52.01	1.43	10
11.00+	54.63	1.89	8	53.20	1.23	11	51.79	1.50	10
11.50+	54.38	2.11	7	53.60	0.75	10	52.35	1.45	8
12.00+	55.07	2.21	9	53.84	1.48	10	52.53	1.51	9
12.50+	54.92	2.06	7	54.15	1.15	11	52.39	1.44	12
13.00+	55.93	1.71	8	54.62	1.26	9	53.16	1.43	8
13.50+	55.68	1.85	6	54.37	1.53	10	53.18	1.62	9
14.00+	56.67	1.57	7	54.79	0.97	10	53.14	1.80	8
14.50+	56.38	1.76	6	54.88	1.36	11	53.00	1.91	8

MODELLING HEAD CIRCUMFERENCE FOR CONTROL MALES

The initial work focused on investigating the ability of the model (3.3) to fit a separate curve for each individual of the control males. Table 4.1.4 gives the correlation coefficients, means and standard deviations of the OLS estimates, where the standard error of g1, skewness, is equal to 0.2642 and for g2, kurtosis, is 0.5226.

The average residual standard deviation is 0.21 cm with the range from 0.10 cm to 0.39 cm. The residual mean square error (RMS) ranges from 0.01 cm^2 to 0.16 cm^2 with an average value of 0.05 cm^2. For a further check of the model, a summary of HC residuals by age intervals is displayed in Table 4.1.5. These errors are close to those reported by Roche, Mukherjee and Guo (1986) and considered acceptable.

Table 4.1.4 Correlation, means and standard deviations of the OLS estimates

	Intercept	t	t^2	$\ln(12t+1)$	$(2-t)_+^3$	$(t-10)_+^3$
Mean	39.8240	-0.2509	0.0093	3.4232	-0.5310	0.0045
S.E.	0.3295	0.0528	0.0023	0.1286	0.0396	0.0010
g1	-0.3697	-0.5798	0.2891	0.6580	0.2269	0.0808
g2	0.6438	0.6319	0.0269	1.0602	0.4525	-0.1976
Correlations						
Intercept	1.0000					
t	0.7873	1.0000				
t^2	-0.7132	-0.9687	1.0000			
$\ln(12t+1)$	-0.8886	-0.9250	0.8371	1.0000		
$(2-t)_+^3$	-0.9327	-0.7592	0.6796	0.8849	1.0000	
$(t-10)_+^3$	0.3664	0.6117	-0.7220	-0.4366	-0.3623	1.0000

Number of individual = 83 (males)

Table 4.1.5 Summary of HC residuals (OLS) by age group for males

Age	n	mean	sd	se	g1	g2
0+	85	0.01	0.17	0.02	-0.23	0.32
0.25+	85	-0.05	0.25	0.03	1.09**	2.06**
0.50+	80	0.00	0.36	0.04	-0.58*	1.63**
0.75+	73	0.10	0.32	0.04	-0.40	0.35
1.00+	82	-0.01	0.30	0.03	0.01	0.43
1.50+	80	-0.09**	0.25	0.03	0.10	0.43
2.00+	81	0.00	0.27	0.03	-0.33	-0.45
2.50+	73	0.01	0.25	0.03	0.33	0.55
3.00+	83	0.02	0.25	0.03	0.13	1.42**
3.50+	82	0.00	0.23	0.03	-0.02	-0.28
4.00+	85	0.03	0.22	0.02	-0.12	-0.03
4.50+	76	0.01	0.21	0.02	-0.35	1.53**
5.00+	83	0.01	0.17	0.02	0.15	0.06
5.50+	78	0.01	0.19	0.02	-0.05	-0.54
6.00+	75	-0.03	0.20	0.02	0.22	-0.05
6.50+	79	-0.01	0.22	0.02	0.08	0.64
7.00+	79	-0.05*	0.19	0.02	0.33	0.52
7.50+	78	0.02	0.19	0.02	0.37	1.07*
8.00+	78	-0.01	0.20	0.02	-0.12	-0.12
8.50+	76	0.00	0.18	0.02	0.17	-0.13
9.00+	84	0.00	0.22	0.02	-0.02	-0.40
9.50+	76	0.00	0.17	0.02	-0.21	0.69
10.00+	83	0.00	0.20	0.02	0.09	0.19
10.50+	78	-0.00	0.23	0.03	-0.01	-0.30
11.00+	80	0.02	0.21	0.02	-0.94**	1.44**
11.50+	78	0.00	0.20	0.02	-0.31	0.42
12.00+	77	0.03	0.16	0.02	0.93**	2.18**
12.50+	79	-0.02	0.20	0.02	0.12	1.14*
13.00+	75	-0.04*	0.21	0.02	0.18	-0.38
13.50+	67	-0.02	0.22	0.03	-0.28	-0.43
14.00+	80	0.03	0.14	0.02	0.46	0.36
14.50+	80	-0.00	0.15	0.02	0.02	-0.27

* P < 0.05; ** P < 0.01

Variance component models of HC for 83 control males using polynomials, conventional splines of equation (2.18) and extended splines of equation (3.3) are illustrated in Figure 4.1.1 in HC-1(*), HC-2(*) and HC-3(*) respectively. The values in brackets are the numbers of fixed parameters of the model, e.g. $HC-1(4)$ is a cubic polynomial with four fixed parameters to be estimated. Values of -2*loglikelihood (LH) and the estimated level 1 residual variance (σ_e^2) are also listed in Figure 4.1.1. To simplify the expressions for the extended splines described in Chapter 3, the combination of symbols will be used in the following text or figures to present the model. For example, a cubic polynomial is written as t^0, t^1, t^2, t^3 and a cubic spline as $t^0, t^1, t^2, t^3, (t-\xi)_+^3$, where t^0 denotes intercept.

The models of $HC-1(3), HC-1(4), HC-1(5), HC-1(6)$ and $HC-1(7)$ are polynomials of degree from 3 to 6; the models of $HC-2(5), HC-2(6)$ and $HC-2(7)$ are the conventional splines and $HC-3(4)$, $HC-3(5)$ and $HC-3(6)$ are the extended splines presented in this study. Comparing the values of LH and σ_e^2 for the models with the same number of parameters one can see clearly that the conventional splines fit better than polynomials and the extended splines fit better than the conventional splines. For example, $HC-1(4)$ and $HC-3(4)$, obtained from a quadratic by adding a different extra term, have the same parameters but with quite different values of LH and σ_e^2. Similarly differences can also be found by comparing $HC-1(5)$ with $HC-2(5)$, $HC-2(5)$ with $HC-3(5)$ etc.

Two knots are chosen at 2 and 10 years for model HC according to rules of thumb (Wold, 1974). Table 4.1.6 shows values of LH and σ_e^2 of the model HC-3(6) with the second knot varying from 7 to 12 in term $(t-\xi)_+^3$ when the first knot is fixed at year 2. The lowest value of LH in the model with knot at 10 is not significantly different from that with knot around 10. It implies that the results of the model with knot around the value of 10 are similar. Table 4.1.6 also shows the LH and σ_e^2 of the model HC-3(6) with the first knot varying from year 1.7 to 2.6 when the second knot is fixed at year 10.

The model HC-3(6) is the variance components model of the model (3.3). The random coefficient model of it is now given:

$$y_{ij} = \beta_{0j} + \beta_{1j}t_{ij} + \beta_{2j}t_{ij}^2 + \beta_{3j}\ln(12t_{ij} + 1) + \beta_{4j}(2 - t_{ij})_+^3 + \beta_{5j}(t_{ij} - 10)_+^3 + e_{ij}, \tag{4.1}$$

with

$$\beta_{0j} = \gamma_0 + u_{0j},$$

$$\beta_{1j} = \gamma_1 + u_{1j},$$

$$\beta_{2j} = \gamma_2 + u_{2j},$$

$$\beta_{3j} = \gamma_3 + u_{3j},$$

$$\beta_{4j} = \gamma_4 + u_{4j}.$$

The estimates of parameters for the model (4.1) is given in Table 4.1.7. The fixed parameters are significantly different from zero ($P < 0.01$) and they allow us to estimate the mean curve for normal males.

Figure 4.1.1 Variance component models of HC for 83 males

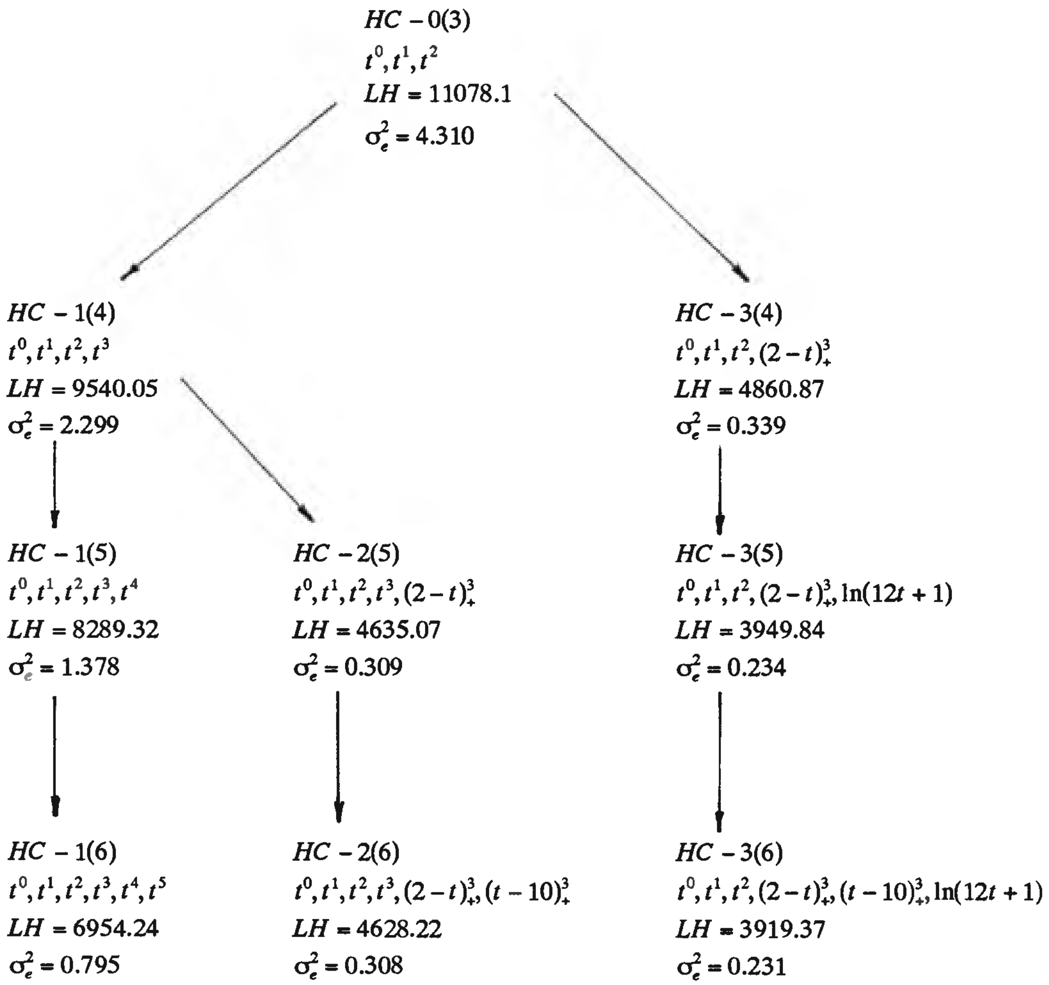


Table 4.1.6 Model $HC-3(6)$ with different knots

	Models	*LH*	σ_e^2
1	$t^0, t^1, t^2, \ln(12t+1), (2-t)_+^3, (t-7)_+^3$	3924.50	0.231
2	$t^0, t^1, t^2, \ln(12t+1), (2-t)_+^3, (t-8)_+^3$	3922.72	0.231
3	$t^0, t^1, t^2, \ln(12t+1), (2-t)_+^3, (t-9)_+^3$	3920.88	0.231
4	$t^0, t^1, t^2, \ln(12t+1), (2-t)_+^3, (t-10)_+^3$	3919.37	0.231
5	$t^0, t^1, t^2, \ln(12t+1), (2-t)_+^3, (t-11)_+^3$	3919.64	0.231
6	$t^0, t^1, t^2, \ln(12t+1), (2-t)_+^3, (t-12)_+^3$	3919.73	0.231
7	$t^0, t^1, t^2, \ln(12t+1), (2.6-t)_+^3, (t-10)_+^3$	3981.71	0.237
8	$t^0, t^1, t^2, \ln(12t+1), (2.5-t)_+^3, (t-10)_+^3$	3969.45	0.236
9	$t^0, t^1, t^2, \ln(12t+1), (2.4-t)_+^3, (t-10)_+^3$	3956.98	0.234
10	$t^0, t^1, t^2, \ln(12t+1), (2.3-t)_+^3, (t-10)_+^3$	3944.76	0.233
11	$t^0, t^1, t^2, \ln(12t+1), (2.2-t)_+^3, (t-10)_+^3$	3933.56	0.232
12	$t^0, t^1, t^2, \ln(12t+1), (2.1-t)_+^3, (t-10)_+^3$	3924.55	0.231
13	$t^0, t^1, t^2, \ln(12t+1), (2.0-t)_+^3, (t-10)_+^3$	3919.37	0.231
14	$t^0, t^1, t^2, \ln(12t+1), (1.9-t)_+^3, (t-10)_+^3$	3920.04	0.231
15	$t^0, t^1, t^2, \ln(12t+1), (1.8-t)_+^3, (t-10)_+^3$	3928.71	0.232
16	$t^0, t^1, t^2, \ln(12t+1), (1.7-t)_+^3, (t-10)_+^3$	3947.08	0.233

Table 4.1.7 Model (4.1) of Head circumference for 83 males

Fixed coefficient	Estimate	s.e.
Intercept	39.83000	0.32600
t	-0.25000	0.05230
t^2	0.00926	0.00227
ln(12t+1)	3.42100	0.12800
$(2-t)_+^3$	-0.53200	0.03930
$(t-10)_+^3$	0.00459	0.00098

Random parameters

Level 2 Covariance matrix (correlations in brackets)

	Intercept	t	t^2	$\ln(12t+1)$	$(2-t)_+^3$	$(t-10)_+^3$
Intercept	6.7240 (1.00)					
t	0.8470 (0.76)	0.1850 (1.00)				
t^2	-0.0340 (-0.70)	-0.0078 (-0.97)	0.0004 (1.00)			
$\ln(12t+1)$	-2.2800 (-0.86)	-0.4070 (-0.92)	0.0160 (0.84)	1.0530 (1.00)		
$(2-t)_+^3$	-0.7480 (-0.92)	-0.0995 (-0.73)	0.0040 (0.67)	0.2780 (0.86)	0.0993 (1.00)	
$(t-10)_+^3$	0.0081 (0.40)	0.0021 (0.64)	-0.0001 (-0.74)	-0.0038 (-0.48)	-0.0010 (-0.41)	0.0001 (1.00)
S.E. of Var.	1.3620	0.0349	0.00007	0.2080	0.0198	0.00001

Level 1 variance = 0.0611 (0.002)
Number of subjects = 83
Number of measurements = 2587

The skewness and the kurtosis of the estimated standardised level 1 residuals are $g_1 = -0.0828$ (0.0487) and $g_2 = 1.1169$ (0.0973) respectively, showing a symmetry for skewness and a sharp peakedness in kurtosis of the distribution. For a check of the model, the plot of standardised residuals by predicted values is given in Figure 4.1.2 which does not show any obvious trend. The Normal plot of level 1 standardised residuals is displayed in Figure 4.1.3, showing an approximately normal distribution except at the two extremes. Further investigation has been made on these extremes and they come from the children who had a deceleration in HC during age 5 months to 1 year. This variability in HC growth can also found in control females (see Figure 4.1.11). Our finding is consistent with the report of Jaffe, Tal, Tirosh and Tamir (1992).

For a check of level 2 residuals, the Normal plots of the standardised level 2 residuals by Normal equivalent scores of the intercept, $t, t^2, \ln(12t + 1), (2 - t)_+^3$ and $(t - 10)_+^3$ are displayed in Figure 4.1.4, showing approximately Normal distributions. The plots of each pair of standardised level 2 residuals are shown in Figures 4.1.5-4.1.7. The likelihood ratio test of the model in table 4.1.7 with its corresponding variance component model reveals that level 2 random variables are significantly different from zero ($\chi^2 = 2328, p < 0.001$).

Figure 4.1.2
Plot of standardised level 1 residuals by predicted values

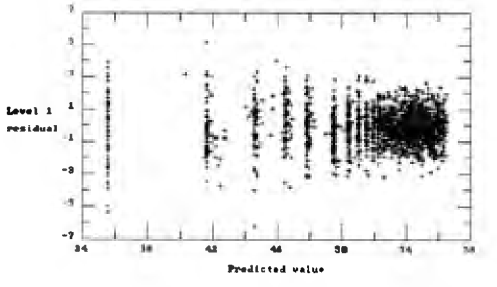


Figure 4.1.3
Plot of standardised level 1 residuals by Normal equivalent scores

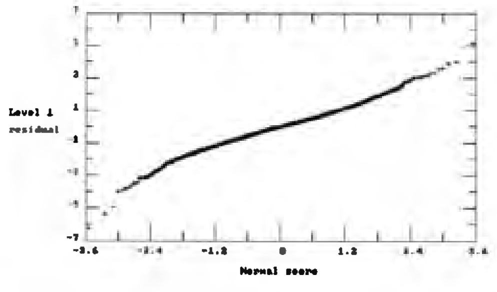
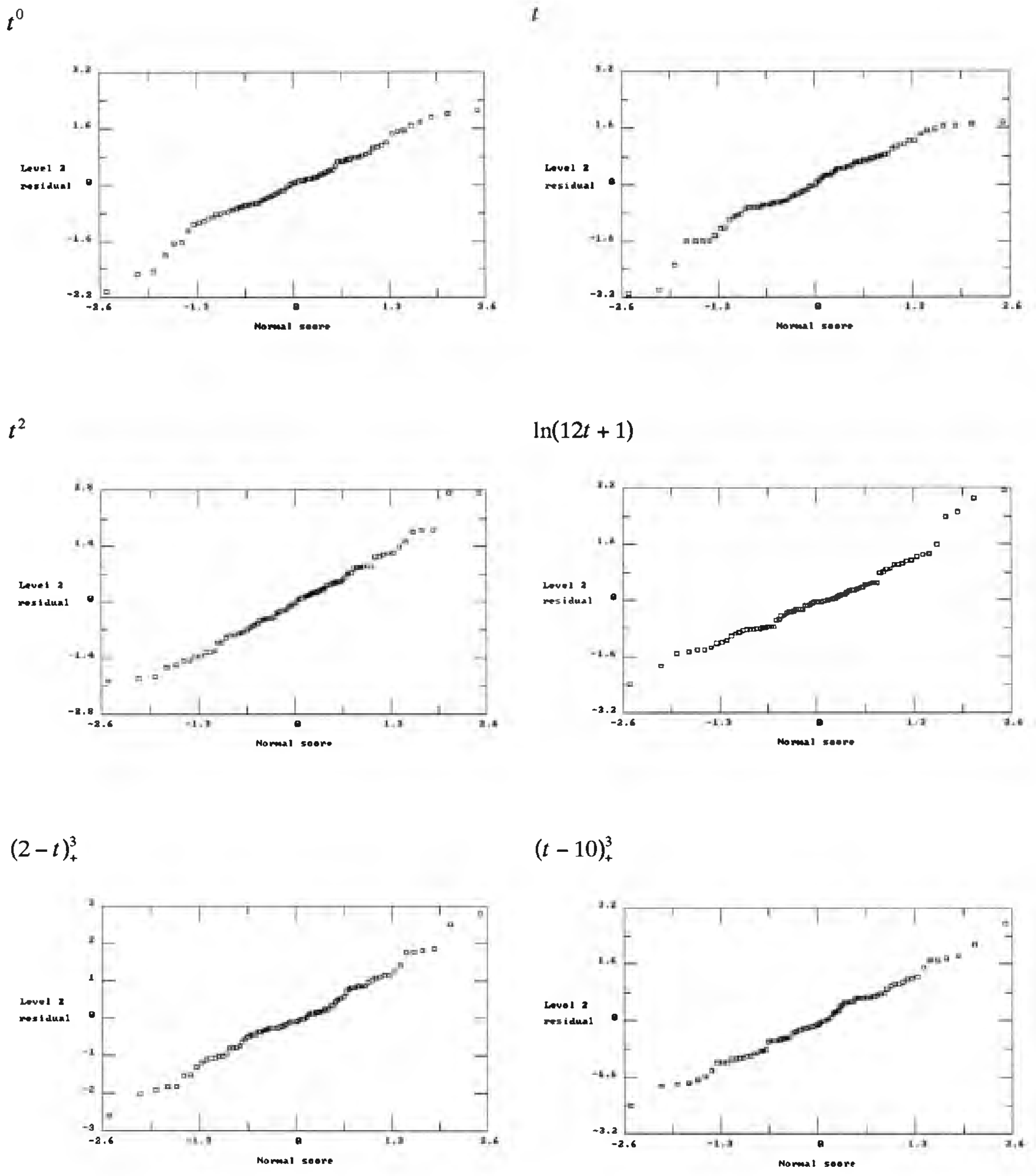


Figure 4.1.4 Standardised level 2 residuals by Normal equivalent scores for the model in Table 4.1.7



t^0

t

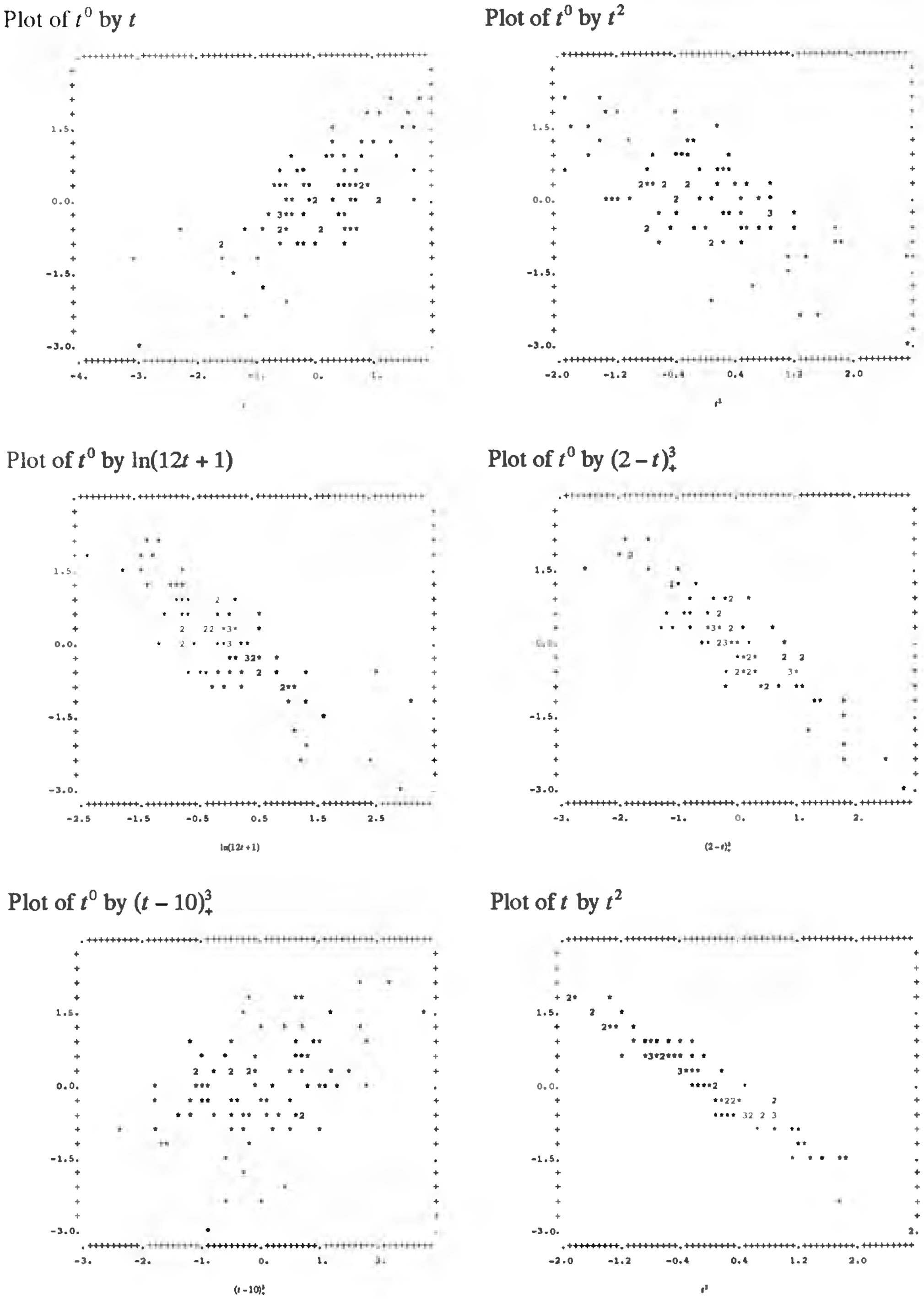
t^2

$\ln(12t+1)$

$(2-t)_+^3$

$(t-10)_+^3$

Figure 4.1.5 Plots of standardised level 2 residuals.



Plot of t^0 by t

Plot of t^0 by t^2

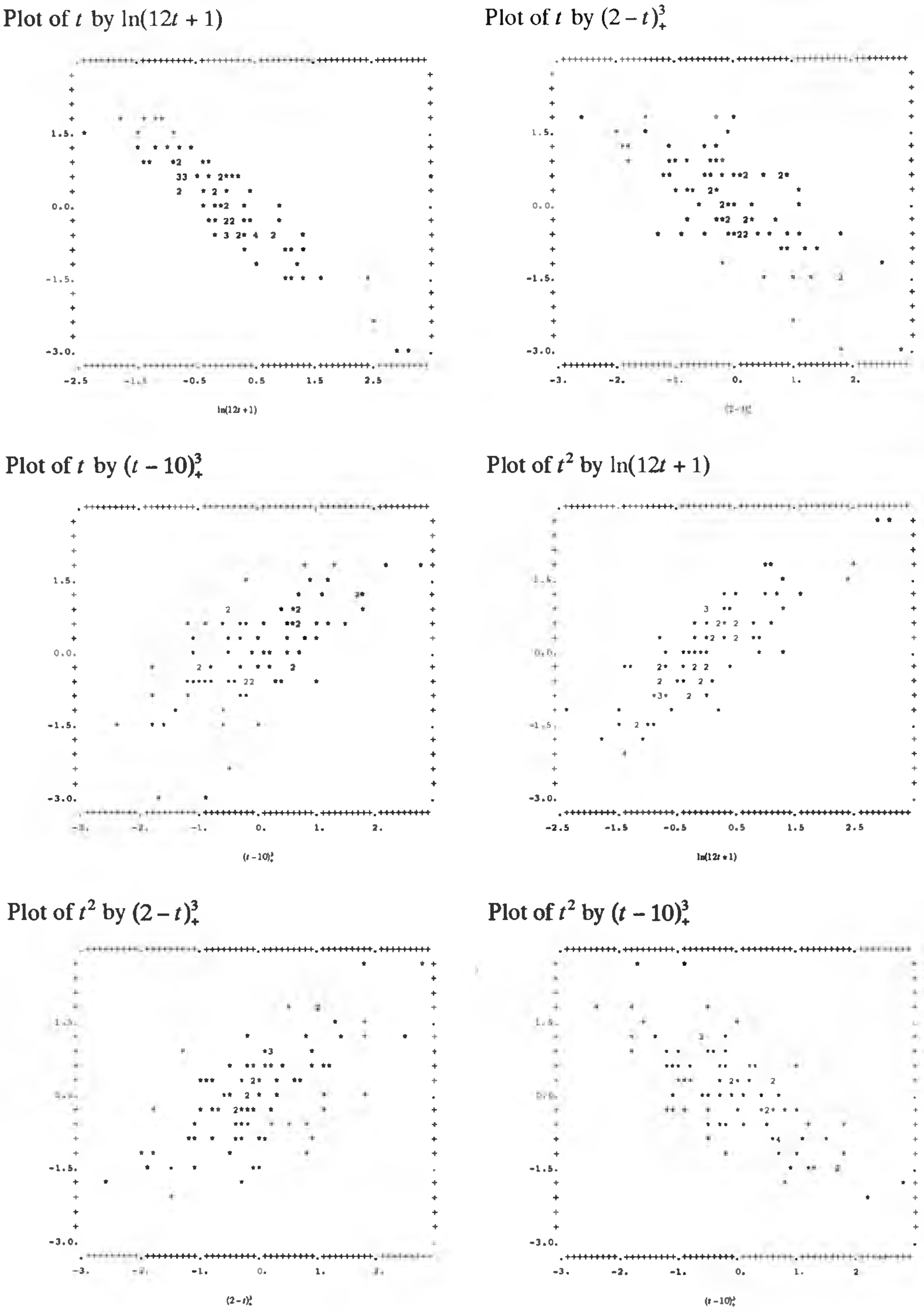
Plot of t^0 by $\ln(12t + 1)$

Plot of t^0 by $(2 - t)^3_+$

Plot of t^0 by $(t - 10)^3_+$

Plot of t by t^2

Figure 4.1.6 Plots of standardised level 2 residuals.



Plot of t by $\ln(12t+1)$

Plot of t by $(2-t)_+^3$

Plot of t by $(t-10)_+^3$

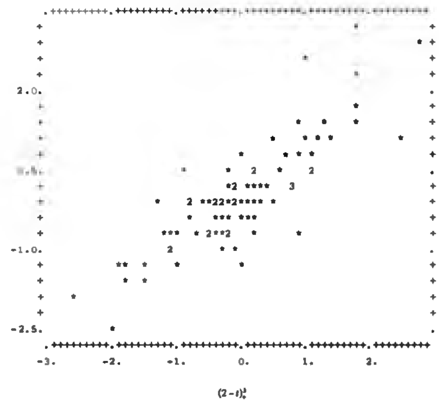
Plot of t^2 by $\ln(12t+1)$

Plot of t^2 by $(2-t)_+^3$

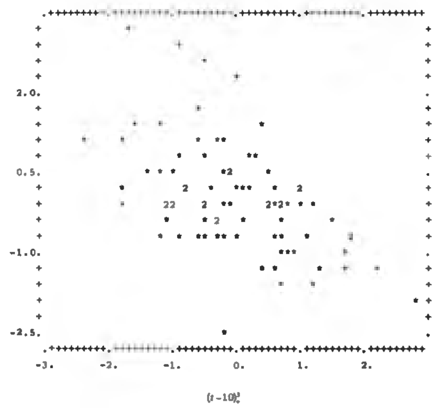
Plot of t^2 by $(t-10)_+^3$

Figure 4.1.7 Plots of standardised level 2 residuals.

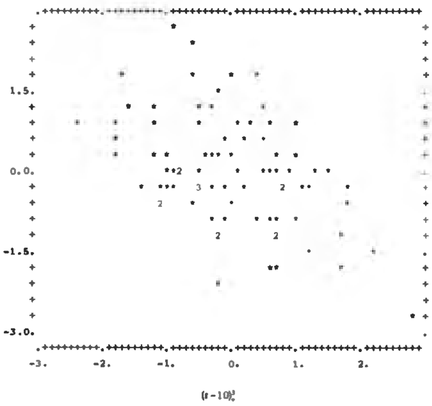
Plot of $\ln(12t + 1)$ by $(2 - t)_+^3$



Plot of $\ln(12t + 1)$ by $(t - 10)_+^3$



Plot of $(2 - t)_+^3$ by $(t - 10)_+^3$



The longitudinally estimated mean population curve of the random coefficient model of Table 4.1.7 with the cross-sectional means plotted is given in Figure 4.1.8. The cross-sectional means are derived from varying numbers of observations within the age intervals. The means estimated by using the multilevel model use the precise age at which the measurement was taken. It is obvious that the multilevel model uses the data more efficiently especially when missing values and unbalanced data are included.

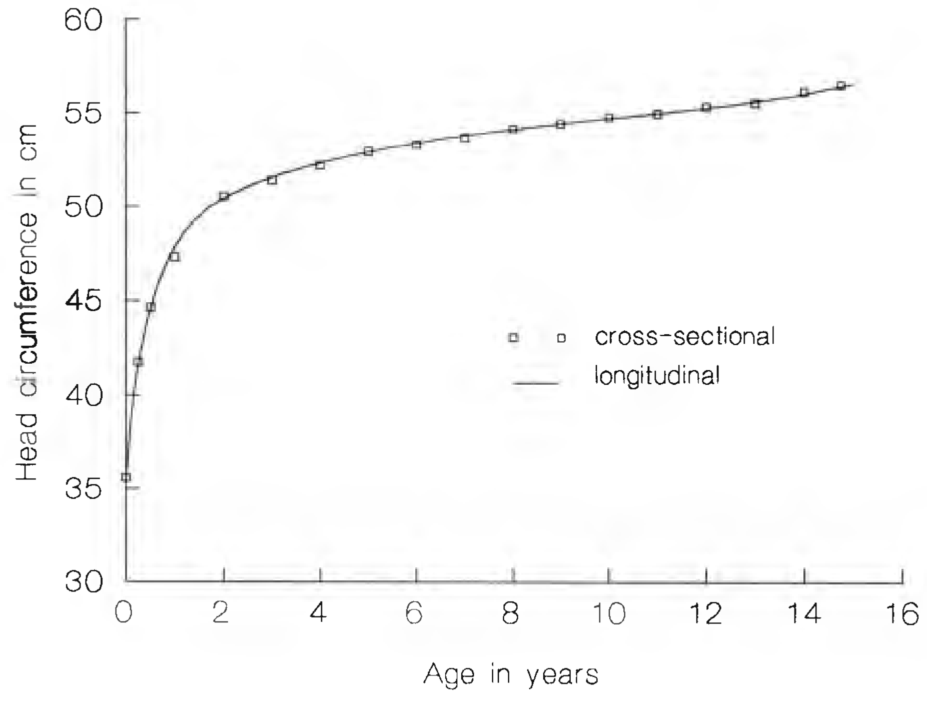


Figure 4.1.8 Longitudinally estimated mean curves (line) of HC for males with the cross-sectional means (dot).

MODELLING HEAD CIRCUMFERENCE FOR CONTROL FEMALES

Firstly ordinary least squares (OLS) for model (3.3) is used to fit the model for individuals of the control females. Table 4.1.8 gives the correlation coefficients, means and standard deviations of the OLS estimates, where the standard error of g1 is equal to 0.3087 and for g2 is 0.6085.

The average residual standard deviation is 0.24 cm with a range from 0.12 cm to 0.44 cm. The residual mean square error (RMS) ranges from 0.02 cm^2 to 0.20 cm^{2}with an average value of 0.06 cm^2. For a further check of the model, a summary of HC residuals by age intervals is displayed in Table 4.1.9. These errors are close to those reported by Roche, Mukherjee and Guo (1986) and considered acceptable.

Table 4.1.8 Correlation, means and standard deviations of the OLS estimates

	Intercept	t	t^2	$\ln(12t+1)$	$(2-t)_+^3$	$(t-10)_+^3$
Mean	39.2220	-0.1434	0.0080	3.1498	-0.5604	-0.0012
S.E.	0.3501	0.0593	0.0027	0.1426	0.0412	0.0013
g1	-0.1726	-0.1783	0.0963	0.3114	0.1406	-0.2376
g2	-0.4396	-0.7204	-0.8407	-0.1998	-0.3528	-0.0267
Correlations						
Intercept	1.0000					
t	0.7569	1.0000				
t^2	-0.6447	-0.9591	1.0000			
$\ln(12t+1)$	-0.8975	-0.9052	0.7863	1.0000		
$(2-t)_+^3$	-0.9530	-0.7809	0.6580	0.9225	1.0000	
$(t-10)_+^3$	0.4400	0.6574	-0.7931	-0.4825	-0.4005	1.0000

Number of individual = 60 (females)

Table 4.1.9 Summary of HC residuals (OLS) by age group for females

Age	n	mean	sd	se	g1	g2
0+	61	0.06**	0.18	0.02	-0.05	0.68
0.25+	58	-0.12**	0.22	0.03	0.20	0.16
0.50+	57	0.00	0.39	0.05	-0.90**	2.98**
0.75+	52	0.05	0.36	0.05	0.07	-0.20
1.00+	58	0.01	0.28	0.04	0.38	0.33
1.50+	58	-0.21**	0.28	0.04	0.31	-0.16
2.00+	56	0.02	0.28	0.04	-0.47	0.01
2.50+	52	0.05	0.26	0.04	0.01	-0.20
3.00+	57	0.10*	0.30	0.04	-0.30	0.67
3.50+	59	0.11**	0.21	0.03	0.69*	1.33*
4.00+	55	0.07*	0.21	0.03	0.41	-0.12
4.50+	64	0.03	0.17	0.02	-0.43	-0.18
5.00+	55	0.01	0.18	0.02	0.19	-0.39
5.50+	54	-0.04	0.20	0.03	-0.21	0.68
6.00+	62	-0.01	0.23	0.03	0.65*	2.06*
6.50+	52	-0.03	0.18	0.02	0.20	-0.53
7.00+	57	-0.01	0.18	0.02	0.36	0.18
7.50+	55	-0.03	0.23	0.03	-0.17	-0.08
8.00+	57	-0.05	0.19	0.03	-0.28	0.51
8.50+	52	-0.03	0.23	0.03	0.19	-0.39
9.00+	62	0.00	0.23	0.03	0.03	0.14
9.50+	54	-0.04	0.20	0.03	-0.37	0.40
10.00+	55	0.01	0.22	0.03	-0.00	0.06
10.50+	55	-0.06*	0.23	0.03	-0.20	-0.08
11.00+	56	-0.01	0.23	0.03	-0.19	0.54
11.50+	58	0.04	0.28	0.04	0.29	0.07
12.00+	55	0.10*	0.26	0.04	0.33	0.18
12.50+	49	0.05	0.26	0.04	0.63	0.51
13.00+	55	-0.06*	0.26	0.03	0.56	0.27
13.50+	56	-0.01	0.26	0.04	-0.24	-0.24
14.00+	53	0.02	0.19	0.03	0.95**	3.28**
14.50+	51	-0.02	0.19	0.03	0.47	-0.06

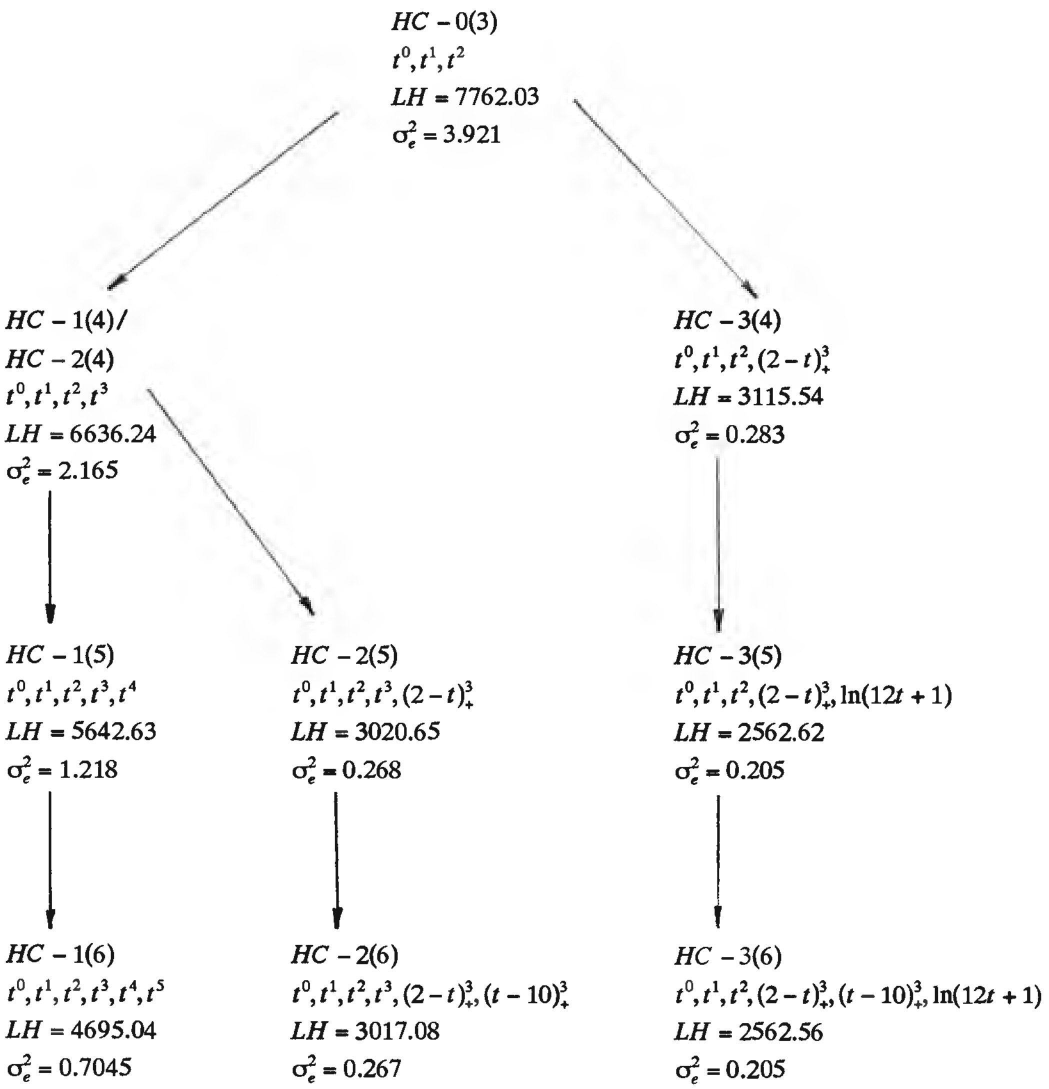
* P < 0.05; ** P < 0.01

Variance component models of HC for 60 control females using polynomials, conventional splines of equation (2.18) and extended splines of equation (3.3) are illustrated in Figure 4.1.9 in HC-1(*), HC-2(*) and HC-3(*) respectively. Similarly to males, it is obvious that the extended splines fit the data better than the polynomials and conventional splines. Table 4.1.10 shows values of LH and σ_e^2 for the model HC-3(6) with the first knot varying around age year 2 and the second knot varying from 7 to 13 in the term $(t - \xi)_+^3$. The lowest value of LH of the model with knot at 12 is not significantly different from that with knots around 10.

Same as for males, the random coefficient model (4.1) is used to the data for females. The estimates of parameters for the model (4.1) are given in Table 4.1.11. All the fixed parameters except $(10 - t)_+^3$ are significant (P < 0.01). It implies that females might not have such an obvious spurt as males have in HC during puberty, which can be seen in Figures 4.1.8 and 4.1.15. The estimated standardised level 1 residuals have skewness, $g_1 = 0.0669$ (0.0579), and kurtosis, $g_2 = 1.7406$ (0.1157), showing a little skewness and a sharp peakedness in kurtosis of the distribution. For a check of the model, the plot of standardised residuals by predicted values is given in Figure 4.1.10 which does not show any obvious trend. The Normal plot of standardised residual is displayed in Figure 4.1.11, showing an approximately Normal distribution except at the two extremes.

For a check of level 2 residuals, the Normal plots of the standardised level 2 residuals by Normal equivalent scores of the $t^0, t, t^2, \ln(12t + 1)$ and $(2 - t)_+^3$ are displayed in Figure 4.1.12, showing an approximately Normal distribution. The plots of each pair of standardised level 2 residuals are given in Figures 4.13-4.14. The likelihood ratio test of the model in Table 4.1.11 with its corresponding variance component model shows that the level 2 random variables are significant (($\chi^2 = 1091.61, p < 0.001$).

Figure 4.1.9 Variance component models of HC for 60 control females



$HC-0(3)$
t^0,t^1,t^2
$LH=7762.03$
$\sigma_e^2=3.921$

$HC-1(4)/$
$HC-2(4)$
t^0,t^1,t^2,t^3
$LH=6636.24$
$\sigma_e^2=2.165$

$HC-3(4)$
$t^0,t^1,t^2,(2-t)_+^3$
$LH=3115.54$
$\sigma_e^2=0.283$

$HC-1(5)$
t^0,t^1,t^2,t^3,t^4
$LH=5642.63$
$\sigma_e^2=1.218$

$HC-2(5)$
$t^0,t^1,t^2,t^3,(2-t)_+^3$
$LH=3020.65$
$\sigma_e^2=0.268$

$HC-3(5)$
$t^0,t^1,t^2,(2-t)_+^3,\ln(12t+1)$
$LH=2562.62$
$\sigma_e^2=0.205$

$HC-1(6)$
t^0,t^1,t^2,t^3,t^4,t^5
$LH=4695.04$
$\sigma_e^2=0.7045$

$HC-2(6)$
$t^0,t^1,t^2,t^3,(2-t)_+^3,(t-10)_+^3$
$LH=3017.08$
$\sigma_e^2=0.267$

$HC-3(6)$
$t^0,t^1,t^2,(2-t)_+^3,(t-10)_+^3,\ln(12t+1)$
$LH=2562.56$
$\sigma_e^2=0.205$

Table 4.1.10 Model $HC-3(6)$ with different knots for females

	Models	LH	σ_e^2
1	$t^0, t^1, t^2, \ln(12t+1), (2-t)_+^3, (t-7)_+^3$	2562.50	0.2050
2	$t^0, t^1, t^2, \ln(12t+1), (2-t)_+^3, (t-8)_+^3,$	2562.60	0.2050
3	$t^0, t^1, t^2, \ln(12t+1), (2-t)_+^3, (t-9)_+^3,$	2562.61	0.2050
4	$t^0, t^1, t^2, \ln(12t+1), (2-t)_+^3, (t-10)_+^3,$	2562.56	0.2050
5	$t^0, t^1, t^2, \ln(12t+1), (2-t)_+^3, (t-11)_+^3,$	2562.51	0.2050
6	$t^0, t^1, t^2, \ln(12t+1), (2-t)_+^3, (t-12)_+^3,$	2562.56	0.2050
7	$t^0, t^1, t^2, \ln(12t+1), (2-t)_+^3, (t-13)_+^3,$	2562.62	0.2050
8	$t^0, t^1, t^2, \ln(12t+1), (2.6-t)_+^3, (t-10)_+^3,$	2591.24	0.2087
9	$t^0, t^1, t^2, \ln(12t+1), (2.5-t)_+^3, (t-10)_+^3,$	2583.09	0.2077
10	$t^0, t^1, t^2, \ln(12t+1), (2.4-t)_+^3, (t-10)_+^3,$	2575.38	0.2068
11	$t^0, t^1, t^2, \ln(12t+1), (2.3-t)_+^3, (t-10)_+^3,$	2568.60	0.2059
12	$t^0, t^1, t^2, \ln(12t+1), (2.2-t)_+^3, (t-10)_+^3,$	2563.48	0.2053
13	$t^0, t^1, t^2, \ln(12t+1), (2.1-t)_+^3, (t-10)_+^3,$	2561.03	0.2051
14	$t^0, t^1, t^2, \ln(12t+1), (2.0-t)_+^3, (t-10)_+^3,$	2562.56	0.2052
15	$t^0, t^1, t^2, \ln(12t+1), (1.9-t)_+^3, (t-10)_+^3,$	2569.53	0.2061
16	$t^0, t^1, t^2, \ln(12t+1), (1.8-t)_+^3, (t-10)_+^3,$	2583.33	0.2077
17	$t^0, t^1, t^2, \ln(12t+1), (1.7-t)_+^3, (t-10)_+^3,$	2604.83	0.2103

Figure 4.1.10
Plot of standardised level 1 residuals by predicted values

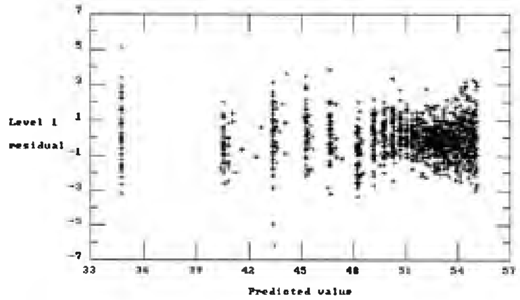


Figure 4.1.11
Plot of standardised level 1 residuals by Normal equivalent scores

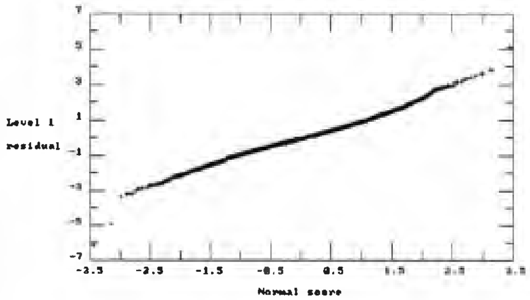


Table 4.1.11 Model (4.1) of head circumference for 60 females

Fixed coefficient	Estimate	s.e.
Intercept	39.25100	0.31970
t	-0.13669	0.04840
t^2	0.00757	0.00193
ln(12t+1)	3.13830	0.12870
$(2-t)_+^3$	-0.56390	0.03852
$(t-10)_+^3$	-0.00074	0.00067

Random parameters

Level 2 Covariance matrix (correlations in brackets)

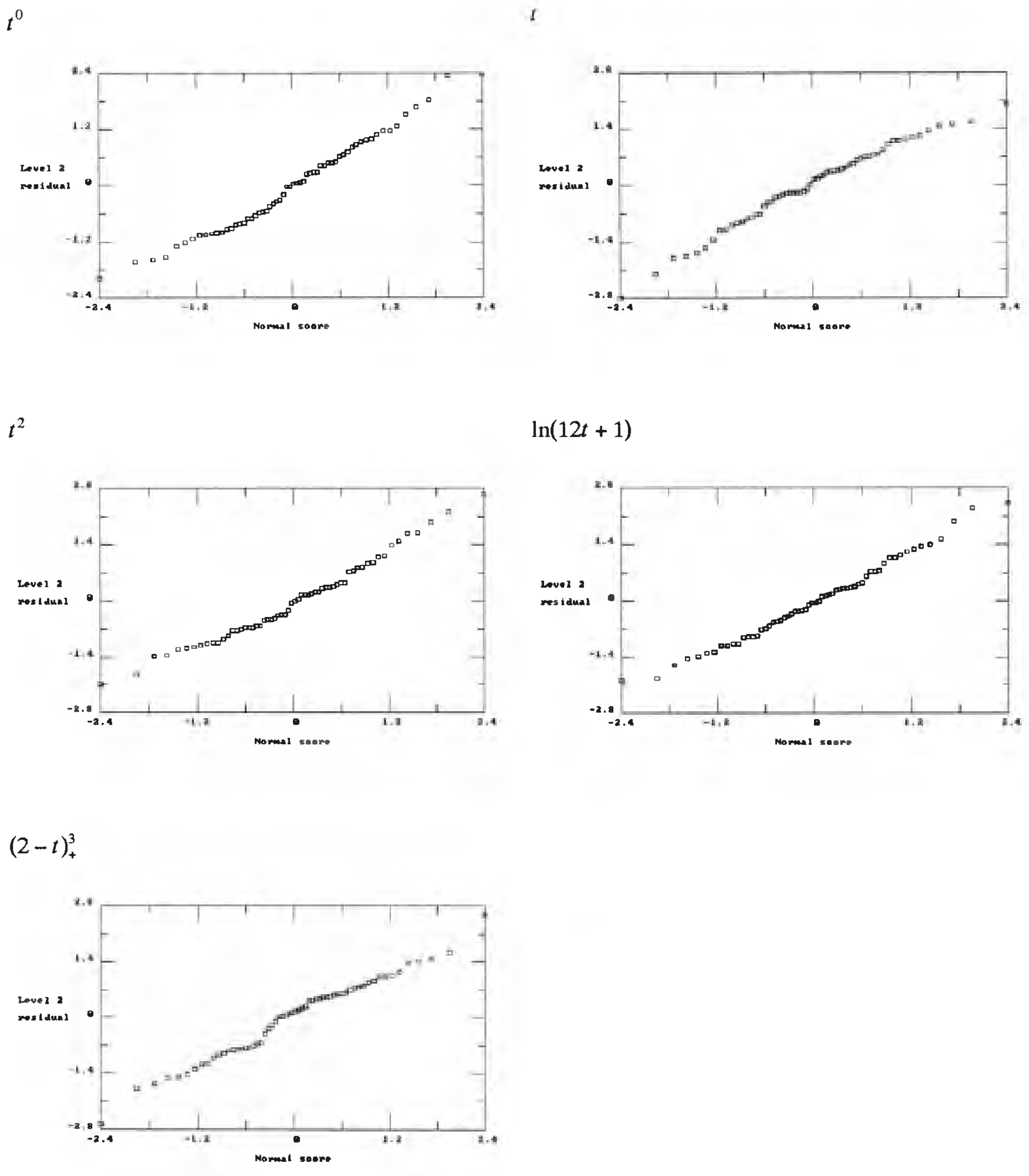
	Intercept	t	t^2	$\ln(12t+1)$	$(2-t)_+^3$
Intercept	3.1203 (1.00)				
t	0.2802 (0.56)	0.0814 (1.00)			
t^2	-0.0072 (-0.38)	-0.0029 (-0.95)	0.0001 (1.00)		
$\ln(12t+1)$	-1.0435 (-0.78)	-0.1845 (-0.86)	0.0056 (0.70)	0.5702 (1.00)	
$(2-t)_+^3$	-0.3509 (-0.91)	-0.0417 (-0.67)	0.0012 (0.51)	0.1427 (0.86)	0.0478 (1.00)
S.E. of Var.	1.0850	0.0228	0.00003	0.1742	0.0159

Level 1 variance = 0.0857 (0.0031)

Number of subjects = 60

Number of measurements = 1789

Figure 4.1.12 Standardised level 2 residuals by Normal equivalent scores for the model in Table 4.1.11.



t^0

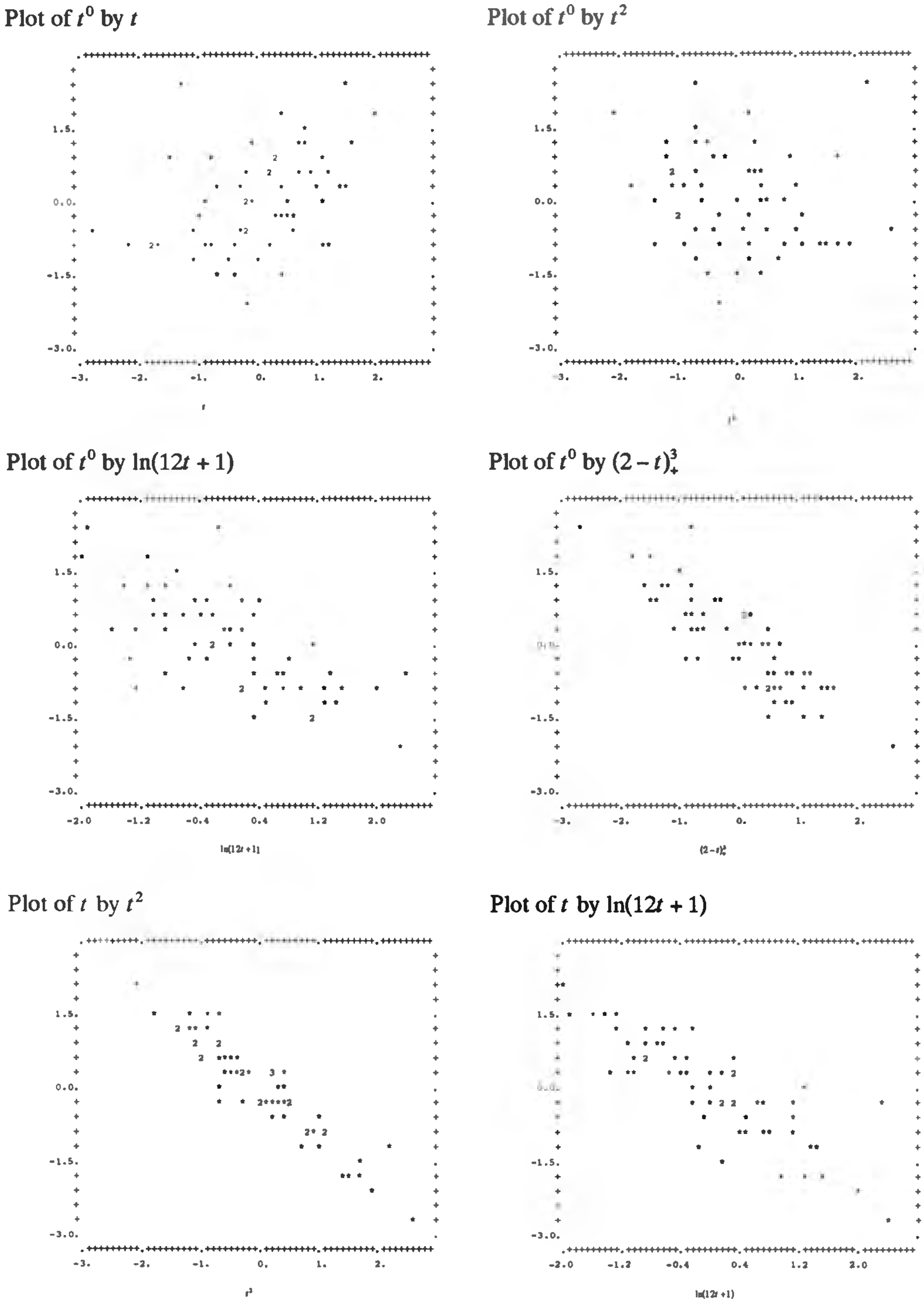
t

t^2

$\ln(12t+1)$

$(2-t)_+^3$

Figure 4.1.13 Plots of standardised level 2 residuals.



Plot of t^0 by t

Plot of t^0 by t^2

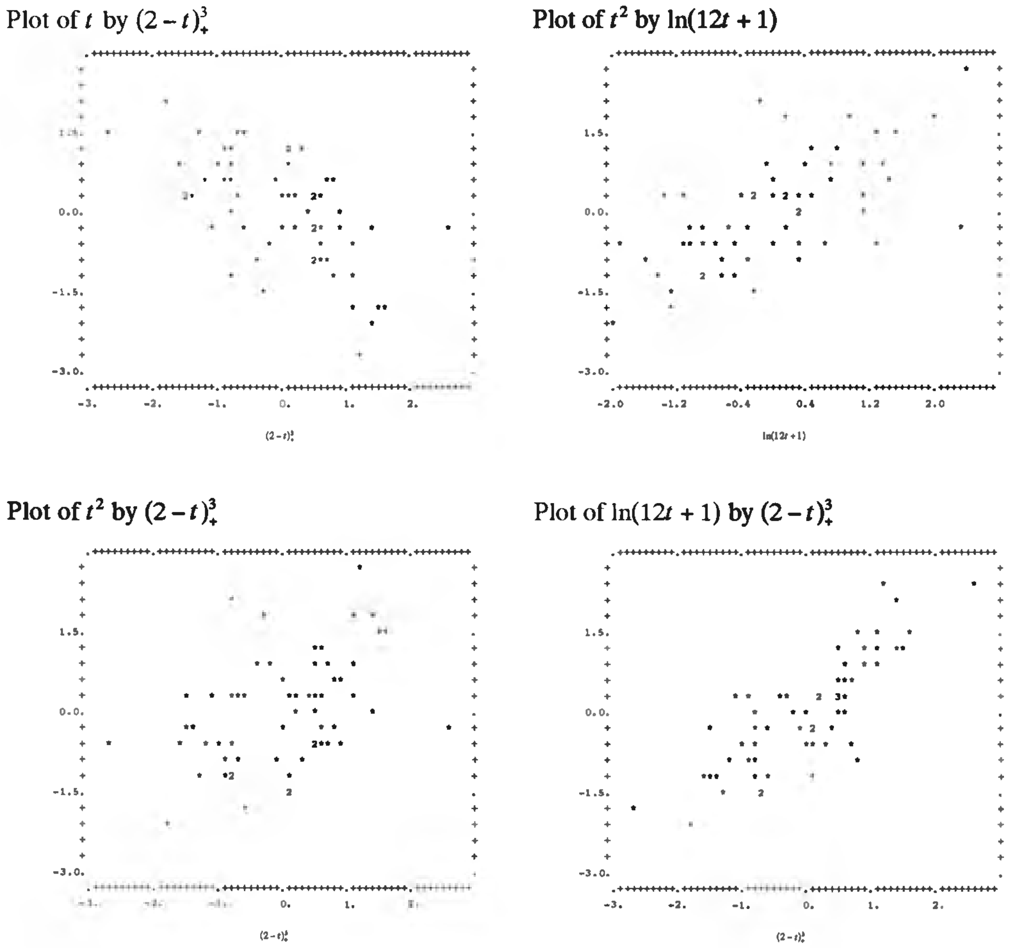
Plot of t^0 by $\ln(12t + 1)$

Plot of t^0 by $(2 - t)_+^3$

Plot of t by t^2

Plot of t by $\ln(12t + 1)$

Figure 4.1.14 Plots of standardised level 2 residuals.



Plot of t by $(2-t)_+^3$

Plot of t^2 by $\ln(12t+1)$

Plot of t^2 by $(2-t)_+^3$

Plot of $\ln(12t+1)$ by $(2-t)_+^3$

The longitudinally estimated mean population curve for the random coefficient model of Table 4.1.11 with the cross-sectional means is shown in Figure 4.1.15. The two sets of means are consistent.

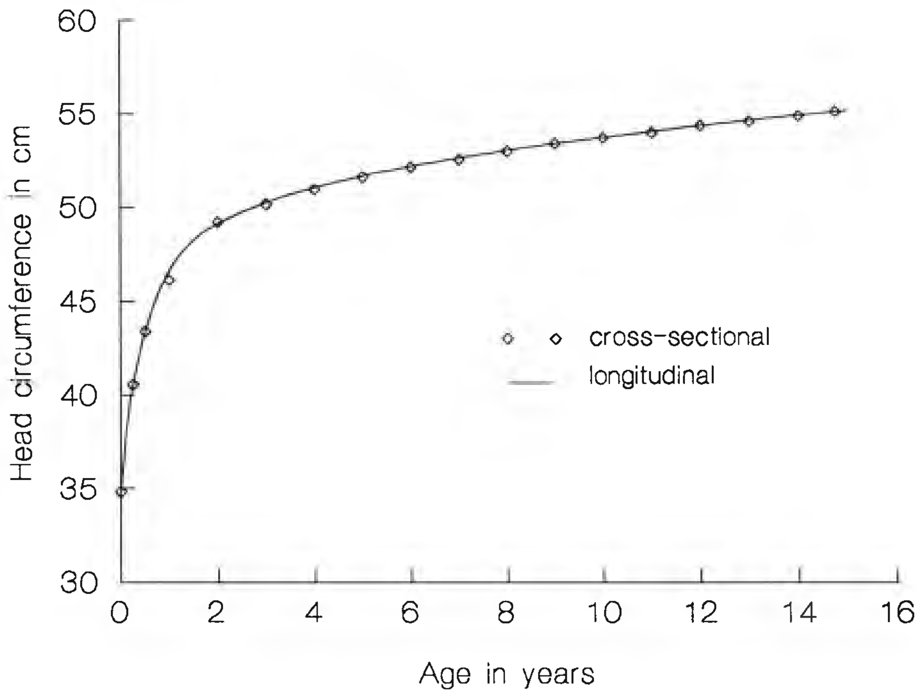


Figure 4.1.15 Longitudinally estimated mean curve (line) of HC for females with the cross-sectional means (dot).

MODELLING HEAD CIRCUMFERENCE WITH KARYOTYPE AS COVARIATE

Modelling head circumference with the covariate of karyotype (XY, XX, XYY, XXY and XXX) is investigated using model HC-3(6). XY indicates normal males, XX normal females, XYY and XXY chromosomally abnormal males, and XXX chromosomally abnormal females.

Firstly the covariate of karyotype is considered as an explanatory variable. That is, in additional to the intercept there are four dummy variables are included into the model HC-3(6): XX-XY, indicating XX versus XY; XYY-XY, indicating XYY versus XY; XXY-XY for XXY versus XY and XXX-XY for XXX versus XY. Table 4.1.12 shows the results of the two-level random coefficient model: the coefficient of XX-XY is estimated as -0.9398 , which means that normal female group on average has a smaller mean than that of the normal male group by 0.94 cm ($p < 0.01$); the mean HC of XXX group is smaller than that of XY by 2.22 cm ($p < 0.01$). For the male group, on average, mean of HC for karyotype XXY is smaller than that of XY by 1.32 cm ($p < 0.01$); mean HC of XYY is only 0.43 cm smaller than that of XY which is not statistically significant ($p > 0.05$) and the mean HC of XXY is about 0.88 cm smaller than that of XYY ($p < 0.05$). The χ^2 values for these tests can be found in table 4.1.13.

The likelihood ratio test is used to test the model of Table 4.1.12 with its corresponding variance component model ($\chi^2 = 4168.6, p < 0.001$).

In the model of Table 4.1.12, the average effect only of covariate on mean HC has been investigated. However, in fact, these differences may change with age.

Table 4.1.12 Model HC-3(6) of head circumference with covariate of karyotype

Fixed coefficient	Estimate	s.e.
Intercept	39.95000	0.25620
t	-0.19780	0.03977
t^2	0.00808	0.00173
ln(12t+1)	3.30700	0.09783
$(2-t)_+^3$	-0.53970	0.02958
$(t-10)_+^3$	0.00287	0.00074
XX-XY	-0.93980	0.17170
XYY-XY	-0.43450	0.33920
XXY-XY	-1.31900	0.32530
XXX-XY	-2.21900	0.33920

Random parameters

Level 2 Covariance matrix (correlations in brackets)

	Intercept	t	t^2	$\ln(12t+1)$	$(2-t)_+^3$	$(t-10)_+^3$
Intercept	7.86600					
	(1.00)					
t	1.06900	0.22720				
	(0.80)	(1.00)				
t^2	-0.04235	-0.00957	0.00043			
	(-0.73)	(-0.97)	(1.00)			
$\ln(12t+1)$	-2.86500	-0.50930	0.02010	1.32200		
	(-0.89)	(-0.93)	(0.84)	(1.00)		
$(2-t)_+^3$	-0.89740	-0.13090	0.00515	0.35850	0.11880	
	(-0.93)	(-0.80)	(0.72)	(0.90)	(1.00)	
$(t-10)_+^2$	0.01054	0.00246	-0.00013	-0.00460	-0.00129	0.00007
	(0.45)	(0.61)	(-0.73)	(-0.47)	(-0.44)	(1.00)
S.E. of Var.	1.11900	0.03003	0.00006	0.18170	0.01652	0.00001

Level 1 variance = 0.06814 (0.01500)

Number of subjects = 174

Number of measurements = 5178

Table 4.1.13 χ^2 values for examining fixed coefficient contrasts

Variable	χ^2	P
XX-XY	29.96	< 0.001
XYY-XY	1.64	> 0.050
XXY-XY	16.45	< 0.001
XXX-XY	42.81	< 0.001

Further investigation of the effects of karyotype and the comparison of mean-parameter curves of the five populations can be found in Table 4.1.14-4.1.17. These results are obtained by modelling all the coefficients of model HC-3(6) to be functions of the four dummy variables of karyotype, XX_XY, XYY_XY, XXY_XY and XXX-XY. Table 4.1.14 shows the estimated fixed parameters for the random coefficient model, in which AGE*XX denotes the term for the interaction of age and XX-XY and is the product of term age and XX-XY; AGE2*XX for the interaction of age^2 and XX-XY; LN*XX for the interaction of $\ln(12t + 1)$ and XX-XY; $(2 - t)_+^3$*XX for the interaction of $(2 - t)_+^3$ and XX-XY; and $(t - 10)_+^3$*XX for the interaction of $(t - 10)_+^3$ and XX-XY. The other interaction terms are similar. Table 4.1.15 shows the random parameters of the random coefficient model of Table 4.1.14. The mean parameters for XY and the difference between other karyotype groups are listed in Table 4.1.16 and the χ^2 values for simultaneous tests and individual tests of the hypotheses are listed in Table 4.1.17.

Table 4.1.14 Model (4.1) of HC with covariate of karyotype (fixed part)

Fixed coefficient	Estimate	s.e.
Intercept	39.83000	0.34410
t	-0.24970	0.05633
t^2	0.00926	0.00247
ln(12t+1)	3.42000	0.13710
$(2-t)_+^3$	-0.53190	0.04188
$(t-10)_+^3$	0.00459	0.00102
XX_XY	-0.61640	0.53210
XYY-XY	-2.27400	1.08000
XXY-XY	-1.33400	1.01000
XXX-XY	-0.38630	1.06800
AGE*XX	0.11030	0.08704
AGE*XYY	-0.16410	0.17570
AGE*XXY	0.01684	0.16530
AGE*XXX	0.37740	0.17430
AGE2*XX	-0.00157	0.00382
AGE2*XYY	0.00188	0.00771
AGE2*XXY	0.00238	0.00727
AGE2*XXX	-0.01523	0.00764
LN*XX	-0.27150	0.21190
LN*XYY	0.74950	0.42920
LN*XXY	-0.09531	0.40230
LN*XXX	-0.95490	0.42520
$(2-t)_+^3$*XX	-0.02785	0.06476
$(2-t)_+^3$*XYY	0.23680	0.13170
$(2-t)_+^3$*XXY	0.01156	0.12290
$(2-t)_+^3$*XXX	-0.20600	0.13020
$(t-10)_+^3$*XX	-0.00548	0.00157
$(t-10)_+^3$*XYY	0.00106	0.00328
$(t-10)_+^3$*XXY	-0.00093	0.00302
$(t-10)_+^3$*XXX	0.00284	0.00317

Table 4.1.15 Model (4.1) of HC with Covariate of Karyotype (random part)

Random parameters

Level 2 Covariance matrix (correlations in brackets)

	Intercept	t	t^2	$\ln(12t+1)$	$(2-t)_+^3$	$(t-10)_+^3$
Intercept	7.47800					
	(1.00)					
t	1.00900	0.21690				
	(0.77)	(1.00)				
t^2	-0.04067	-0.00928	0.00042			
	(-0.71)	(-0.97)	(1.00)			
$\ln(12t+1)$	-2.67800	-0.47890	0.01928	1.22900		
	(-0.87)	(-0.92)	(0.84)	(1.00)		
$(2-t)_+^3$	-0.85200	-0.12410	0.00496	0.33690	0.11340	
	(-0.92)	(-0.77)	(0.70)	(0.89)	(1.00)	
$(t-10)_+^2$	0.01085	0.00256	-0.00013	-0.00486	-0.00132	0.00006
	(0.47)	(-0.68)	(-0.78)	(-0.53)	(-0.47)	(1.00)
S.E. of Var.	1.07500	0.02885	0.00006	0.17120	0.01591	0.00001

Level 1 variance = 0.06812 (0.0015)

Number of subjects = 174

Number of measurements = 5178

In Table 4.1.16, the column XY shows the mean estimated parameters for the normal male group with the standard error in brackets, and the column XX-XY the gender difference between the normal female and male groups. The column XYY-XY shows parameters of the difference between the chromosomally abnormal male XYY and the normal male group, similarly for the other karyotypes.

Simultaneous tests of hypotheses about parameter contrasts are made and the χ^2 values of these tests are listed in Table 4.1.17. In the female group, the parameters of the mean population curves are significantly different between the control XX and the chromosomally abnormal XXX ($p < 0.001$). In the male group, the parameters of the mean population curves are significantly different between the control XY and chromosomally abnormal XXY ($p < 0.001$), however the sets of parameters of XY and XYY curves are not significantly different ($p > 0.05$). There is a significant difference of the mean population curves between the normal XX and XY ($p < 0.001$). These results are consistent with those reported by Ratcliffe, Masera, Pan and McKie (1994).

The likelihood ratio test is used for the model in Table 4.1.15 with its corresponding variance component model and again shows a significant difference($p < 0.001$). The distributional assumptions are checked by the level 1 and level 2 residuals. The estimated standardised level 1 residuals have skewness $g_1 = -0.0044$ (0.0340), and kurtosis, $g_2 = 2.0631$ (0.0680), which show symmetry and a sharp peakedness of the distribution. The plot of standardised residuals by predicted values is given in Figure 4.1.16, and does not show any obvious trend. The Normal plot of standardised level 1 residuals by Normal equivalent scores is displayed in Figure 4.1.17. The Normal plots of the standardised level 2 residuals are displayed in Figure 4.1.18 and show an approximately Normal distribution. Figures 4.1.19-4.1.21 are the plots of the standardised level 2 residuals plotted against each other.

Table 4.1.16 Mean parameters (S.E.) and the karyotype effects

Parameter	XY**	XX-XY**	XYY-XY	XXY-XY**	XXX-XY**
CONS	39.83000	-0.61640	-2.27400	-1.33400	-0.38630
	(0.34410)	(0.53210)	(1.08000)	(1.01000)	(1.06800)
t	-0.24970	0.11030	-0.16410	0.01684	0.37740
	(0.05633)	(0.08704)	(0.17570)	(0.16530)	(0.17430)
t^2	0.00926	-0.00157	0.00188	0.02381	-0.01523
	(0.00247)	(0.00382)	(0.00771)	(0.00727)	(0.00764)
$\ln(12t+1)$	3.42000	-0.27150	0.74950	-0.09531	-0.95490
	(0.13710)	(0.21190)	(0.42920)	(0.40230)	(0.42520)
$(2-t)_+^3$	-0.53190	-0.02785	0.23680	0.01156	-0.20600
	(0.04188)	(0.06476)	(0.13170)	(0.12290)	(0.13020)
$(t-10)_+^3$	0.00459	-0.00548	0.00106	-0.00093	0.00284
	(0.00102)	(0.00157)	(0.00328)	(0.00302)	(0.00317)

** $P < 0.01$ of simultaneous tests of the hypotheses by F test.

Table 4.1.17 χ^2 values for examining karyotype effects

Karyotype	χ^2	P
Male Group		
XYY / XY	11.78	0.06706
XXY / XY	22.80	0.00087
XYY / XXY	19.10	0.00390
Female Group		
XXX / XX	23.23	0.00072
Normal Male/Female		
XX / XY	68.54	0.00000

Figure 4.1.16

Plot of standardised level 1 residuals by predicted values

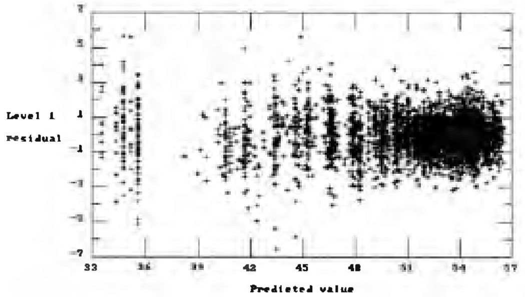


Figure 4.1.17

Plot of Standardised Level 1 Residuals by Normal equivalent scores

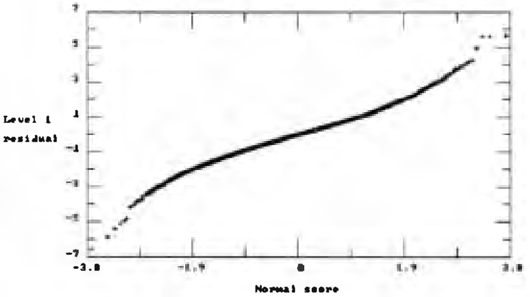
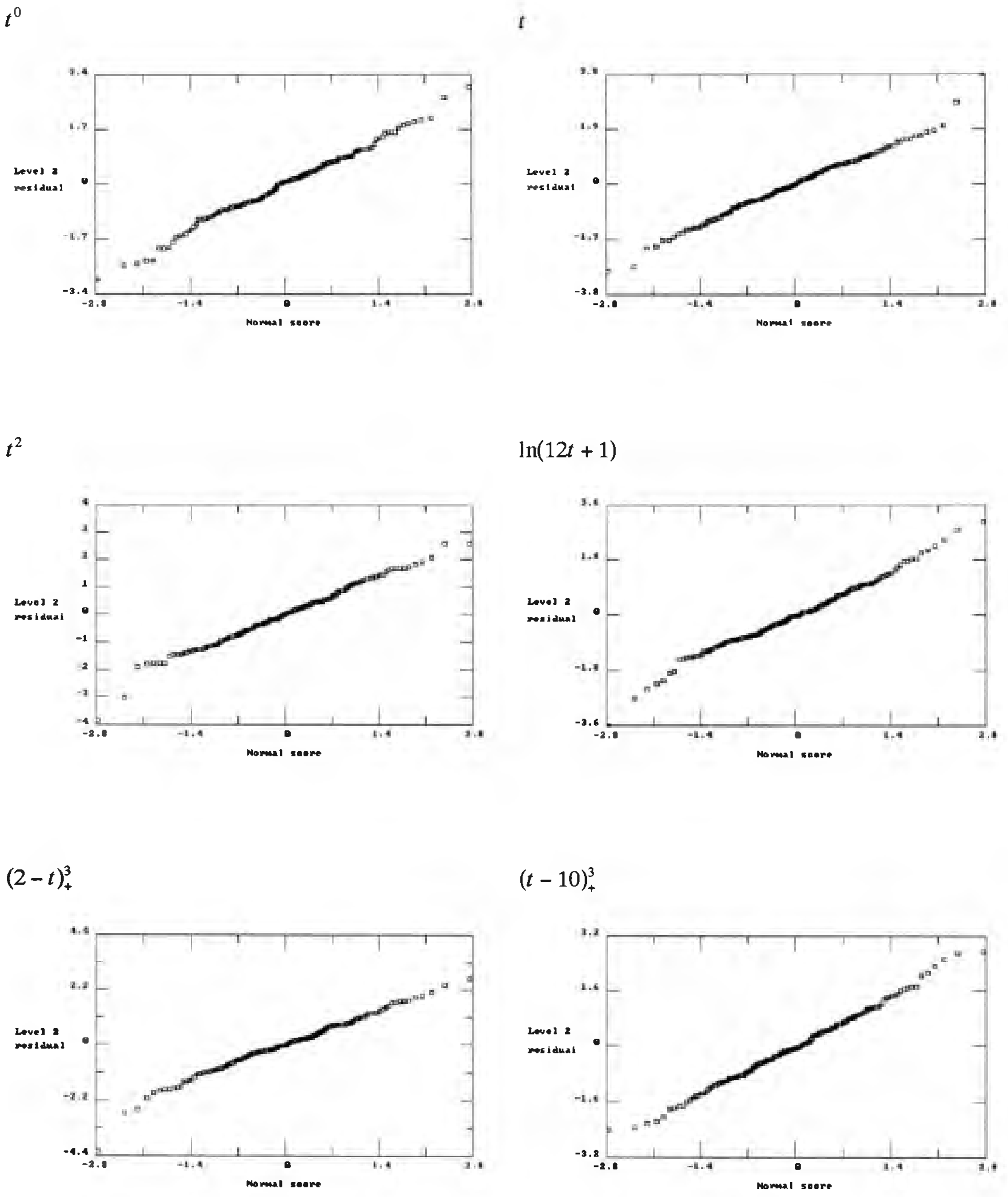


Figure 4.1.18 Standardised level 2 residuals by Normal equivalent scores for the model in Table 4.1.15.



t^0

t

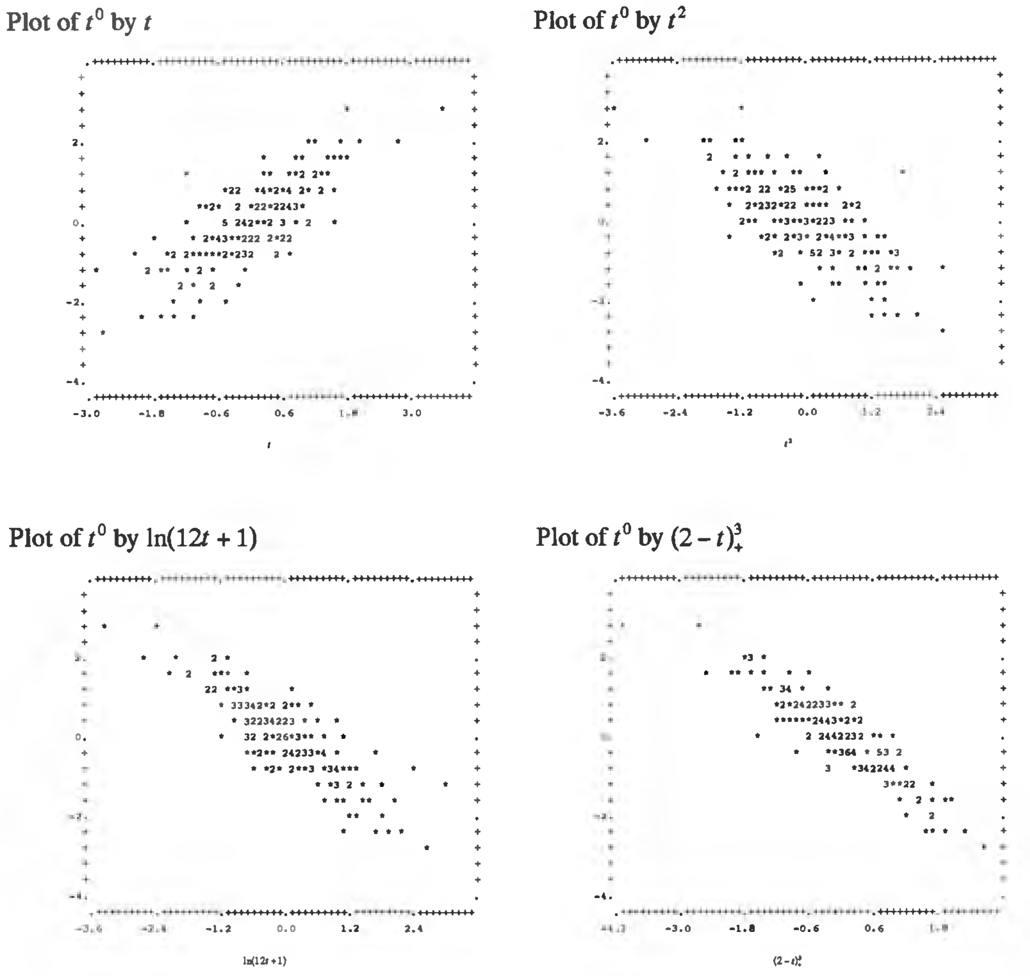
t^2

$\ln(12t+1)$

$(2-t)_+^3$

$(t-10)_+^3$

Figure 4.1.19 Plots of standardised level 2 residuals.



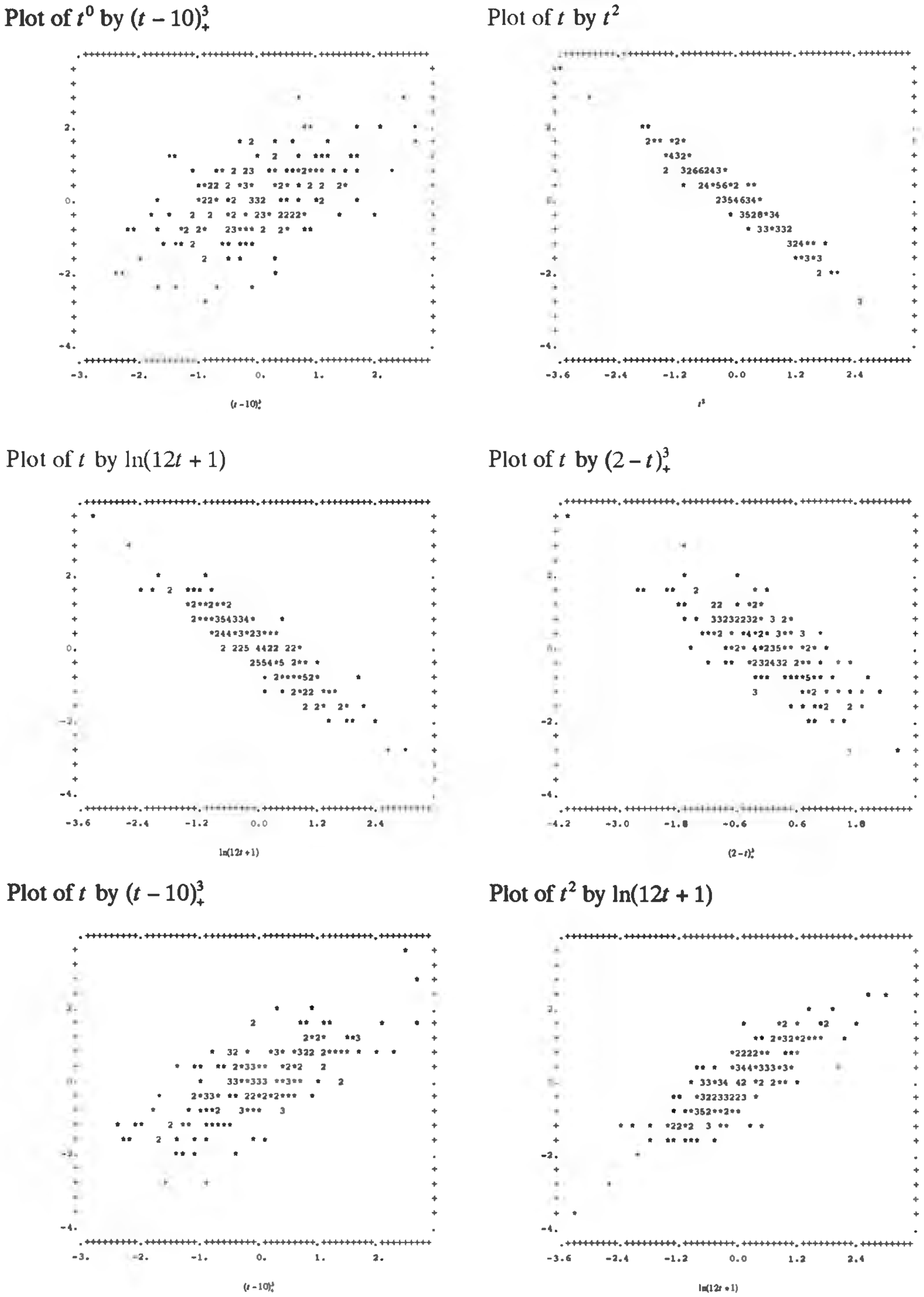
Plot of t^0 by t

Plot of t^0 by t^2

Plot of t^0 by $\ln(12t+1)$

Plot of t^0 by $(2-t)_+^3$

Figure 4.1.20 Plots of standardised level 2 residuals.



Plot of t^0 by $(t-10)_+^3$

Plot of t by t^2

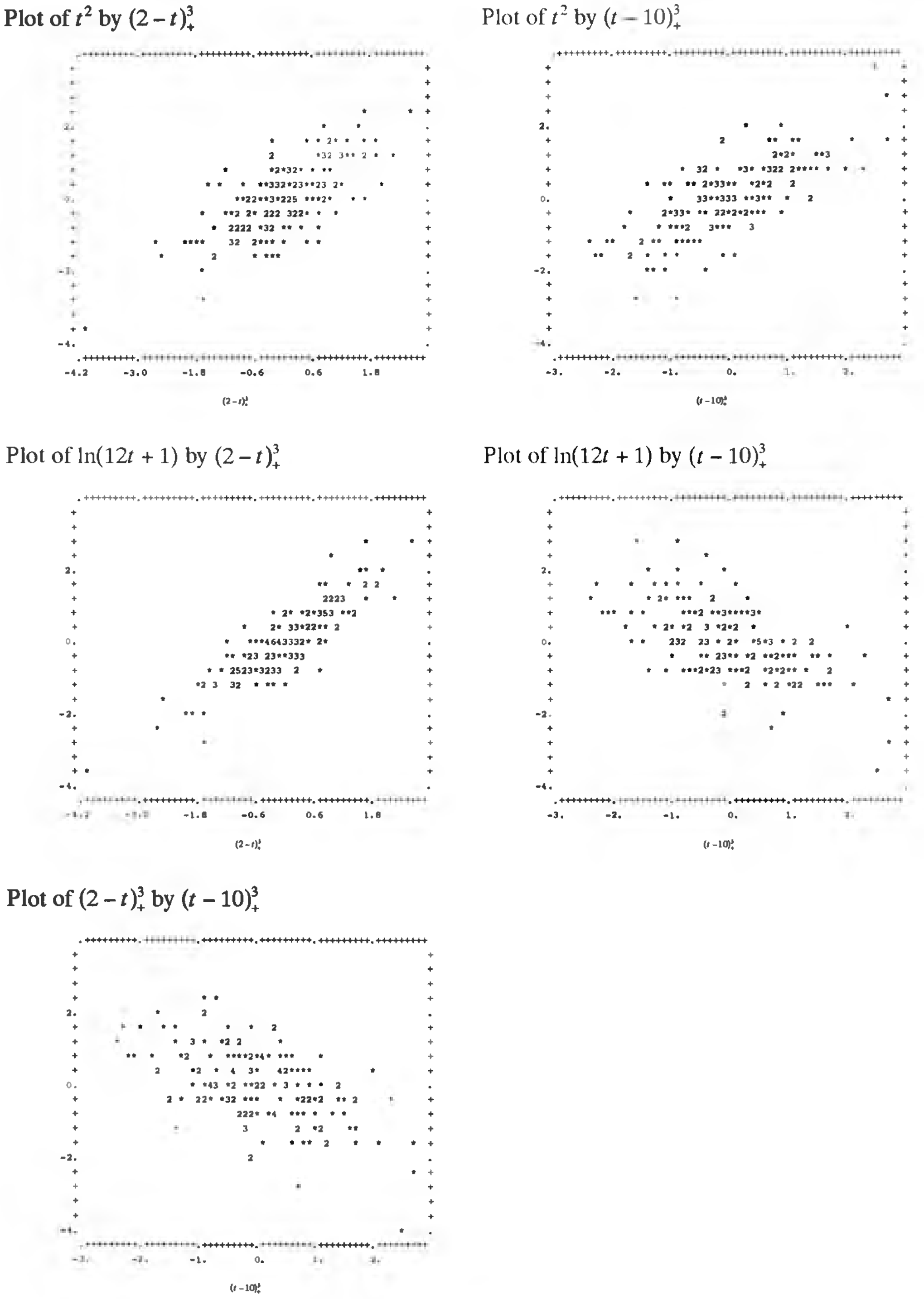
Plot of t by $\ln(12t+1)$

Plot of t by $(2-t)_+^3$

Plot of t by $(t-10)_+^3$

Plot of t^2 by $\ln(12t+1)$

Figure 4.1.21 Plots of standardised level 2 residuals.



Plot of t^2 by $(2-t)_+^3$

Plot of t^2 by $(t-10)_+^3$

Plot of $\ln(12t+1)$ by $(2-t)_+^3$

Plot of $\ln(12t+1)$ by $(t-10)_+^3$

Plot of $(2-t)_+^3$ by $(t-10)_+^3$

The predicted mean values of head circumference for the five populations are presented in Figure 4.1.22, which are consistent with the curves of Ratcliffe, Masera, Pan and McKie (1994).

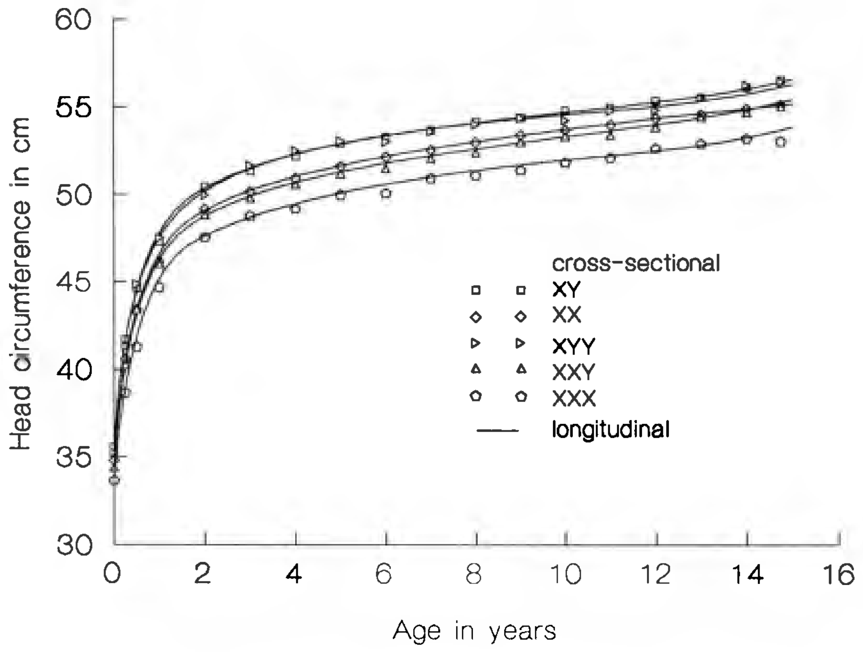


Figure 4.1.22 Longitudinally estimated mean curves of HC for XY, XX, XYY, XXY and XXX (line) with the cross-sectional means (dot).

4.2 Modelling Height (HT)

The subjects studied include 89 males and 67 females from the control group (99 males and 74 females) of the Edinburgh Longitudinal study, initiated in 1972 and they are known to be chromosomally normal as they were born at a time when the Medical Research Council was conducting a newborn cytogenetic survey (Ratcliffe and Paul 1986). These children were at least 16 years of age for males and 15 years for females in 1992 when this study started. Anthropometric measures were taken by S.G.Ratcliffe who was trained in measurement techniques by the late R.H. Whitehouse of the Department of Growth and Development, Institute of Child Health, University of London. The children were measured 3-monthly during the first year of life and twice-yearly thereafter. The data used in this section cover ages from 0.25 to 18.5 years.

Gross errors have been checked. Tables 4.2.1-4.2.2 show the number of measures and the cross-sectional mean for each age group with height in cm and age in years.

Table 4.2.1 The number of measures (HT) by gender

Gender	Individuals	Measurements	Mean measures per individual
Males	89	3044	34.2
Females	67	2134	31.7

Table 4.2.2 Mean height (cm) and number of measures by age group

	XY			XX		
Age Group	Mean	S.D.	N	Mean	S.D.	N
0.25+	62.75	2.60	90	60.83	2.06	62
0.50+	68.62	2.65	84	66.71	2.25	62
0.75+	72.55	2.52	76	70.93	2.18	56
1.00+	76.25	2.81	86	74.33	2.60	61
1.50+	81.68	3.04	82	80.19	2.91	59
2.00+	87.21	3.17	81	85.58	3.15	60
2.50+	91.67	3.47	76	90.27	3.46	54
3.00+	95.89	3.73	90	94.57	3.14	63
3.50+	99.46	3.97	86	98.62	3.24	63
4.00+	103.57	4.15	93	102.04	3.76	61
4.50+	106.40	4.47	82	106.01	3.98	67
5.00+	110.38	4.61	87	109.45	4.03	59
5.50+	113.17	4.71	83	112.93	4.12	58
6.00+	116.58	4.94	82	116.07	4.84	66
6.50+	119.58	5.17	84	117.98	4.17	54
7.00+	122.38	5.03	87	122.19	4.83	65
7.50+	125.60	5.34	82	125.05	5.19	57
8.00+	128.31	5.37	84	127.51	4.84	64
8.50+	131.13	5.82	80	131.08	4.81	57
9.00+	134.08	5.71	90	132.86	5.61	68
9.50+	136.69	5.61	81	136.14	6.01	57
10.00+	139.29	5.94	88	139.13	5.58	62
10.50+	141.29	6.52	82	141.70	6.64	60
11.00+	144.40	5.91	86	145.01	7.37	62
11.50+	146.84	6.68	84	148.32	7.95	65
12.00+	149.97	6.78	80	151.38	7.39	59
12.50+	152.27	7.38	85	153.55	7.92	53
13.00+	155.91	7.61	79	156.86	7.46	59
13.50+	160.01	7.54	73	159.20	7.21	61
14.00+	164.11	8.05	85	160.59	6.69	60
14.50+	167.67	7.86	81	162.09	6.45	49
15.00+	169.89	7.29	79	163.34	6.14	60
15.50+	172.64	7.17	76	164.09	6.15	39
16.00+	174.39	7.06	81	163.74	6.40	64
16.50+	176.46	6.67	72	164.23	6.12	39
17.00+	176.31	6.73	61	164.15	7.35	35
17.50+	176.66	6.93	52	166.51	6.47	21
18.00+	177.41	6.86	33	168.92	5.87	12

MODELLING HEIGHT FOR CONTROL MALES

Firstly the model (3.5) of Model HT-A in section 3 is investigated by fitting curves for each individual using OLS. The knots at years 9, 11, 13, 15, 17.5 are chosen according to rules of thumb (Wold, 1974). The average residual standard deviation is 0.5517 cm with a range from 0.2918 cm to 0.9121 cm. Table 4.2.3 gives the correlation coefficients, means and standard deviations of the OLS estimates, where the standard error of g1 is equal to 0.2554 and of g2 is 0.5056.

The residual mean square error (RMS) ranges from 0.09 cm^2 to 0.83 cm^2 at an average value of 0.32 cm^2. For a further check of the model, the summary of residuals by age intervals is displayed in Table 4.2.4. The results are close to expectation (see Bock, Wainer, Petersen, Thissen, Murray and Roche, 1973; Berkey 1982b).

Variance component models of height for 89 control males using polynomials, conventional splines of equation (2.18) and extended splines of equation (3.5) are illustrated in Figure 4.2.1 in HT-1(*), HT-2(*) and HT-3(*) respectively. The notation used in section 4.2 is same as that in section 4.1.

The models of $HT-1(3)$, $HT-1(4)$, $HT-1(5)$, $HT-1(6)$, $HT-1(7)$ and $HT-1(8)$ are polynomials of order from 3 to 7; the models of $HT-2(5)$, $HT-2(6)$, $HT-2(7)$ and $HT-2(8)$ are the conventional splines and $HT-3(3)$, $HT-3(4)$, $HT-3(5)$, $HT-3(6)$, $HT-3(7)$ and $HT-3(8)$ are the extended splines presented in this study. Comparing the values of LH and σ_e^2 for the models with seven or eight parameters we can see obviously that the extended splines fit the data better than the conventional splines. Comparing HT-3(7) with HT-3(8) we can also see that the term $(9-t)_+^3$ in HT-3(8) is not significantly different from zero ($p > 0.05$), that is, the knot at age 9 is not necessary in the population mean curve for males.

Table 4.2.5 shows values of *LH* and σ_e^2 of these models derived from HT-3(7). The models with last knot at 17 years, such as model 1 and 2 in Table 4.2.5, are not better than HT-3(7). Models 6 to 13 are the model HT-3(8) with first knot varying from age 1 to 8 years and show that they are close to model HT-3(8) and HT-3(7).

In addition the model (3.6) of Model HT-B in section 3 is tested by adding the term of $1/t$ to the model HT-3(7). Comparing model 1 with 2 or 3 with 4 in Table 4.2.5 we can find that $1/t$ is not significantly different from zero (P > 0.05).

The random coefficient model of HT-3(7) is now used for analysing the data for males and is given:

$$y_{ij} = \beta_{0j} + \beta_{1j}t_{ij} + \beta_{2j}\ln(t_{ij}) + +\beta_{3j}(11 - t_{ij})_+^3 + \beta_{4j}(13 - t_{ij})_+^3 + \beta_{5j}(15 - t_{ij})_+^3 + \beta_{6j}(17.5 - t_{ij})_+^3 + e_{ij}, \quad (4.2)$$

with

$$\beta_{0j} = \gamma_0 + u_{0j} + u_{3j}(t_{ij} - 9)_+^2 + u_{4j}(t_{ij} - 11)_+^2,$$

$$\beta_{1j} = \gamma_1 + u_{1j},$$

$$\beta_{2j} = \gamma_2 + u_{2j},$$

$$\beta_{3j} = \gamma_3,$$

$$\beta_{4j} = \gamma_4,$$

$$\beta_{5j} = \gamma_5,$$

$$\beta_{6j} = \gamma_{6.}$$

Table 4.2.6 gives the estimates of parameters for the random coefficient model (4.2). Examining fixed coefficient contrasts shows that all the fixed coefficients are significantly different from zero (P < 0.01).

Table 4.2.3 Correlation, means and standard deviations of the OLS estimates

	t^0	lnt	$t-9$	$(9-t)_+^3$	$(11-t)_+^3$	$(13-t)_+^3$	$(15-t)_+^3$	$(17.5-t)_+^3$
Mean	150.100	6.3553	1.0294	0.0091	0.1551	-0.6131	0.6694	-0.2233
S.E.	2.5538	0.2727	0.3198	0.0183	0.0524	0.0713	0.0520	0.0154
g1	-0.1632	-0.2674	-0.2923	0.6472	-1.1020	0.9357	-0.1608	-0.2361
g2	3.9072	-0.4879	8.1899	-0.0109	0.4975	0.7801	1.9578	5.0345

Correlations

t^0	1.0000							
lnt	-0.1445	1.0000						
$t-9$	-0.9137	-0.1532	1.0000					
$(9-t)_+^3$	0.2521	0.3078	-0.3752	1.0000				
$(11-t)_+^3$	-0.1529	-0.1249	0.2240	-0.9360	1.0000			
$(13-t)_+^3$	-0.1738	0.0286	0.1355	0.7417	-0.9135	1.0000		
$(15-t)_+^3$	0.5341	0.0357	-0.5261	-0.4289	0.6573	-0.9035	1.0000	
$(17.5-t)_+^3$	-0.7783	-0.0706	0.7934	0.1131	-0.3438	0.6829	-0.9291	1.0000

Number of individual = 89 (males)

Note t^0 denotes intercept.

Table 4.2.4 Summary of HT residuals (OLS) by age group for males

Age	n	mean	sd	se	g1	g2
0.25+	90	0.05	0.34	0.04	0.24	0.29
0.50+	84	0.06	0.56	0.06	0.08	0.68
0.75+	76	-0.03	0.60	0.07	0.18	0.51
1.00+	86	-0.18*	0.62	0.07	0.01	-0.14
1.50+	82	-0.11	0.71	0.08	-0.07	-0.10
2.00+	81	0.00	0.59	0.07	-0.06	0.32
2.50+	76	0.18*	0.60	0.07	0.05	0.68
3.00+	90	0.23**	0.45	0.05	0.40	-0.16
3.50+	86	0.10*	0.44	0.05	0.39	-0.28
4.00+	93	0.00	0.45	0.05	-0.08	0.07
4.50+	82	-0.08*	0.40	0.04	-0.17	-0.19
5.00+	87	-0.13**	0.45	0.05	-0.21	-0.08
5.50+	83	-0.06	0.42	0.05	0.15	-0.67
6.00+	82	-0.19**	0.43	0.05	-0.06	-0.57
6.50+	84	-0.14**	0.40	0.04	-0.26	-0.28
7.00+	87	0.04	0.47	0.05	-0.26	0.64
7.50+	82	0.00	0.43	0.05	0.51	0.20
8.00+	84	0.18**	0.45	0.05	-0.07	-0.02
8.50+	80	0.07	0.42	0.05	0.01	-0.36
9.00+	90	0.09	0.41	0.04	0.16	-0.54
9.50+	81	0.03	0.46	0.05	0.07	-0.11
10.00+	88	0.04	0.51	0.05	0.00	-0.40
10.50+	82	-0.12*	0.48	0.05	0.26	0.54
11.00+	86	-0.09	0.42	0.05	-0.10	0.76
11.50+	84	0.01	0.55	0.06	-0.43	0.60
12.00+	80	-0.04	0.79	0.09	-0.85**	1.73**
12.50+	85	0.03	0.66	0.07	0.13	-0.35
13.00+	79	0.05	0.56	0.06	-0.01	-0.61
13.50+	73	0.03	0.65	0.08	0.00	0.69
14.00+	85	0.07	0.88	0.10	-0.03	0.43
14.50+	85	-0.08	0.81	0.09	-0.60*	0.35
15.00+	79	-0.15*	0.59	0.07	-0.32	1.16*
15.50+	76	-0.06	0.60	0.07	0.22	-0.38
16.00+	81	0.05	0.83	0.09	0.63*	-0.10
16.50+	74	0.18*	0.73	0.09	0.92**	0.72
17.00+	61	0.17**	0.45	0.06	0.63*	0.54
17.50+	52	-0.11	0.67	0.09	-1.16**	1.41*
18.00+	33	-0.23	0.87	0.15	-0.91*	-0.24

* p 0.05; ** P < 0.01

Figure 4.2.1 Variance component models of HT for 83 males

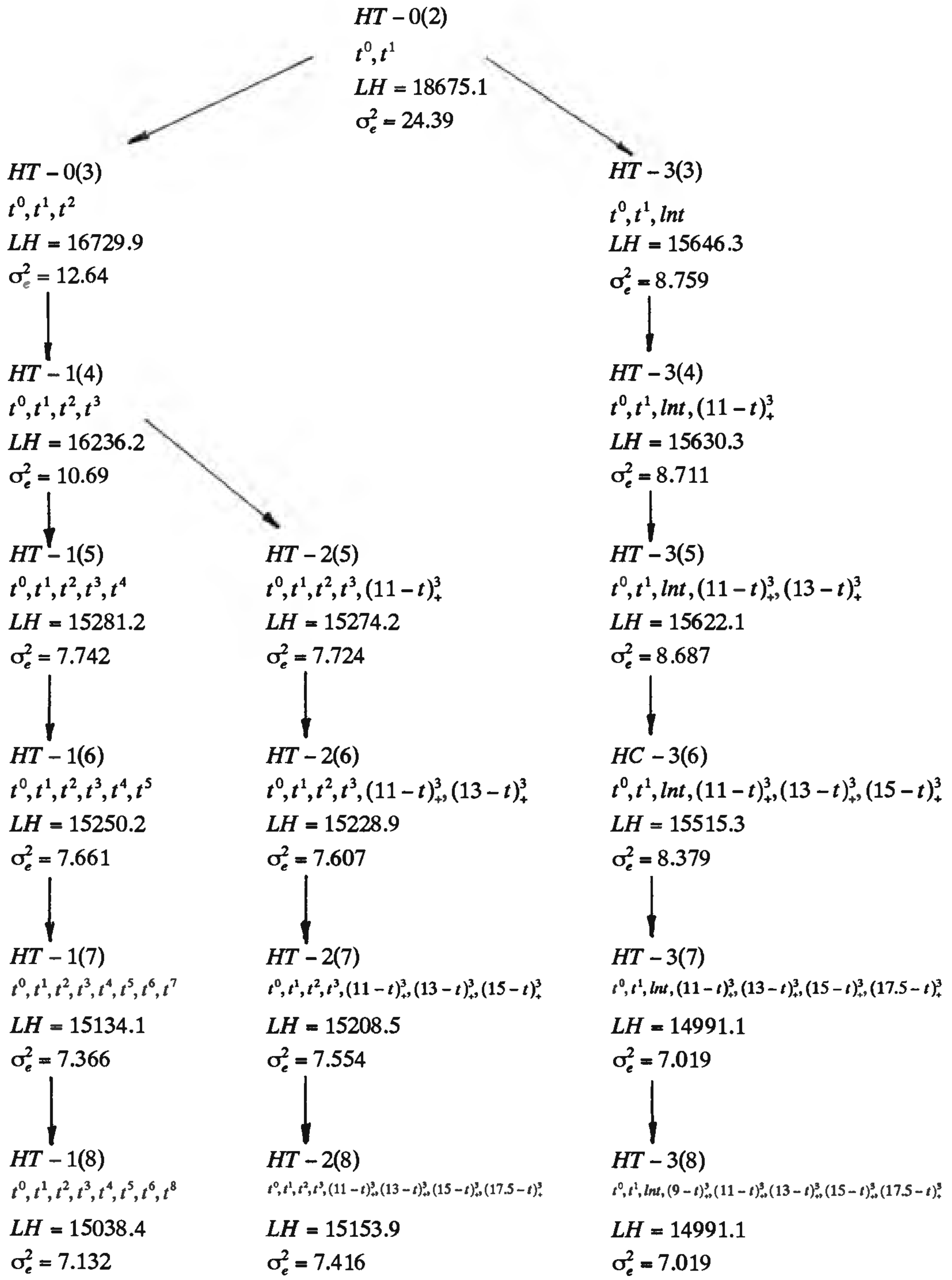


Table 4.2.5 Model $HT-3(7)$ with different knots or terms for males

	Models	LH	σ_e^2
1	$t^0, t^1, lnt, (11-t)_+^3, (13-t)_+^3, (15-t)_+^3, (17-t)_+^3,$	14994.3	7.026
2	$t^0, t^1, lnt, (11-t)_+^3, (13-t)_+^3, (15-t)_+^3, (17-t)_+^3, 1/t$	14994.0	7.026
3	$t^0, t^1, lnt, (11-t)_+^3, (13-t)_+^3, (15-t)_+^3, (17.5-t)_+^3$	14991.1	7.019
4	$t^0, t^1, lnt, (11-t)_+^3, (13-t)_+^3, (15-t)_+^3, (17.5-t)_+^3, 1/t$	14991.0	7.018
5	$t^0, t^1, lnt, (9-t)_+^3, (11-t)_+^3, (13-t)_+^3, (15-t)_+^3, (17.5-t)_+^3,$	14991.1	7.019
6	$t^0, t^1, lnt, (8-t)_+^3, (11-t)_+^3, (13-t)_+^3, (15-t)_+^3, (17.5-t)_+^3$	14991.1	7.019
7	$t^0, t^1, lnt, (7-t)_+^3, (11-t)_+^3, (13-t)_+^3, (15-t)_+^3, (17.5-t)_+^3$	14990.9	7.018
8	$t^0, t^1, lnt, (6-t)_+^3, (11-t)_+^3, (13-t)_+^3, (15-t)_+^3, (17.5-t)_+^3$	14990.5	7.017
9	$t^0, t^1, lnt, (5-t)_+^3, (11-t)_+^3, (13-t)_+^3, (15-t)_+^3, (17.5-t)_+^3$	14990.1	7.016
10	$t^0, t^1, lnt, (4-t)_+^3, (11-t)_+^3, (13-t)_+^3, (15-t)_+^3, (17.5-t)_+^3$	14990.0	7.016
11	$t^0, t^1, lnt, (3-t)_+^3, (11-t)_+^3, (13-t)_+^3, (15-t)_+^3, (17.5-t)_+^3$	14990.7	7.018
12	$t^0, t^1, lnt, (2-t)_+^3, (11-t)_+^3, (13-t)_+^3, (15-t)_+^3, (17.5-t)_+^3$	14990.1	7.016
13	$t^0, t^1, lnt, (1-t)_+^3, (11-t)_+^3, (13-t)_+^3, (15-t)_+^3, (17.5-t)_+^3$	14990.7	7.018

Table 4.2.6 Model (4.2) of height for 89 males

Fixed coefficient	Estimate	s.e.
t^0	148.9900	1.7200
Int	6.2657	0.2195
$t-9$	0.1697	0.0112
$(11-t)_+^3$	-0.6164	0.0171
$(13-t)_+^3$	0.6599	0.0134
$(15-t)_+^3$	-0.2179	0.0057
$(17.5-t)_+^3$	1.2258	0.2101

Random parameters

Level 2 Covariance matrix (correlations in brackets)

	t^0	*Int*	$t-9$	$(t-9)_+^2$	$(t-11)_+^2$
t^0	25.6930				
	(1.00)				
Int	-0.9369	2.0321			
	(-0.13)	(1.00)			
$t-9$	2.2437	-0.3037	0.2470		
	(0.89)	(-0.43)	(1.00)		
$(t-9)_+^2$	-0.3219	0.0814	-0.0504	0.1264	
	(-0.18)	(0.16)	(-0.28)	(1.00)	
$(t-11)_+^2$	0.1019	-0.1380	0.0447	-0.1237	0.3826
	(0.03)	(-0.16)	(0.15)	(-0.97)	(1.00)
S.E. of Var.	4.1290	0.3492	0.0415	0.0202	0.0603

Level 1 variance = 1.0298 (0.0286)

Number of subjects = 89

Number of measurements = 3044

The Likelihood ratio test on the model in Table 4.2.6 compared to its corresponding variance component model is significant ($p < 0.001$). The estimated standardised level 1 residuals' skewness, $g_1 = 0.0620$ (0.0444), and kurtosis, $g_2 = 1.7944$ (0.0887), show symmetry and a sharp peakedness of the distribution. For a check of the model, the plot of standardised residuals by predicted values is given in Figure 4.2.2, in which the residuals are randomly distributed and the spread of the residuals increases with the magnitude of the predicted values. However, we have been unable to model this using a complex level 1 variance-covariance structure. The Normal plot of standardised level 1 residuals is displayed in Figure 4.2.3, showing an approximately Normal distribution.

The terms of $(t-9)_+^2$ and $(t-11)_+^2$ are chosen to join other variables to form the level 2 random part of the model in Table 4.2.6. For a check of level 2 residuals, the Normal plots of the standardised level 2 residuals by Normal equivalent scores are displayed in Figure 4.2.4 and the plots of each pair of standardised level 2 residuals are shown in Figures 4.2.5-4.2.6.

Higher-order terms, such as cubic, $(9-t)_+^3$ etc. should be useful in order to explain between-individual variation and to reduce level 1 variance. Unfortunately with our limited number of level 2 units it is not possible to obtain estimates.

Figure 4.2.2
Plot of standardised level 1 residual by
predicted values

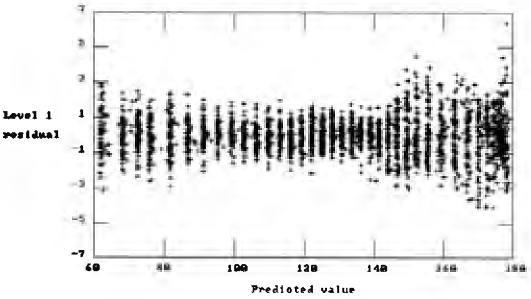


Figure 4.2.3
Plot of standardised level 1 residuals by
Normal equivalent scores

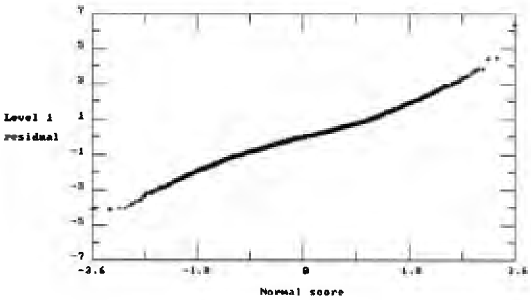
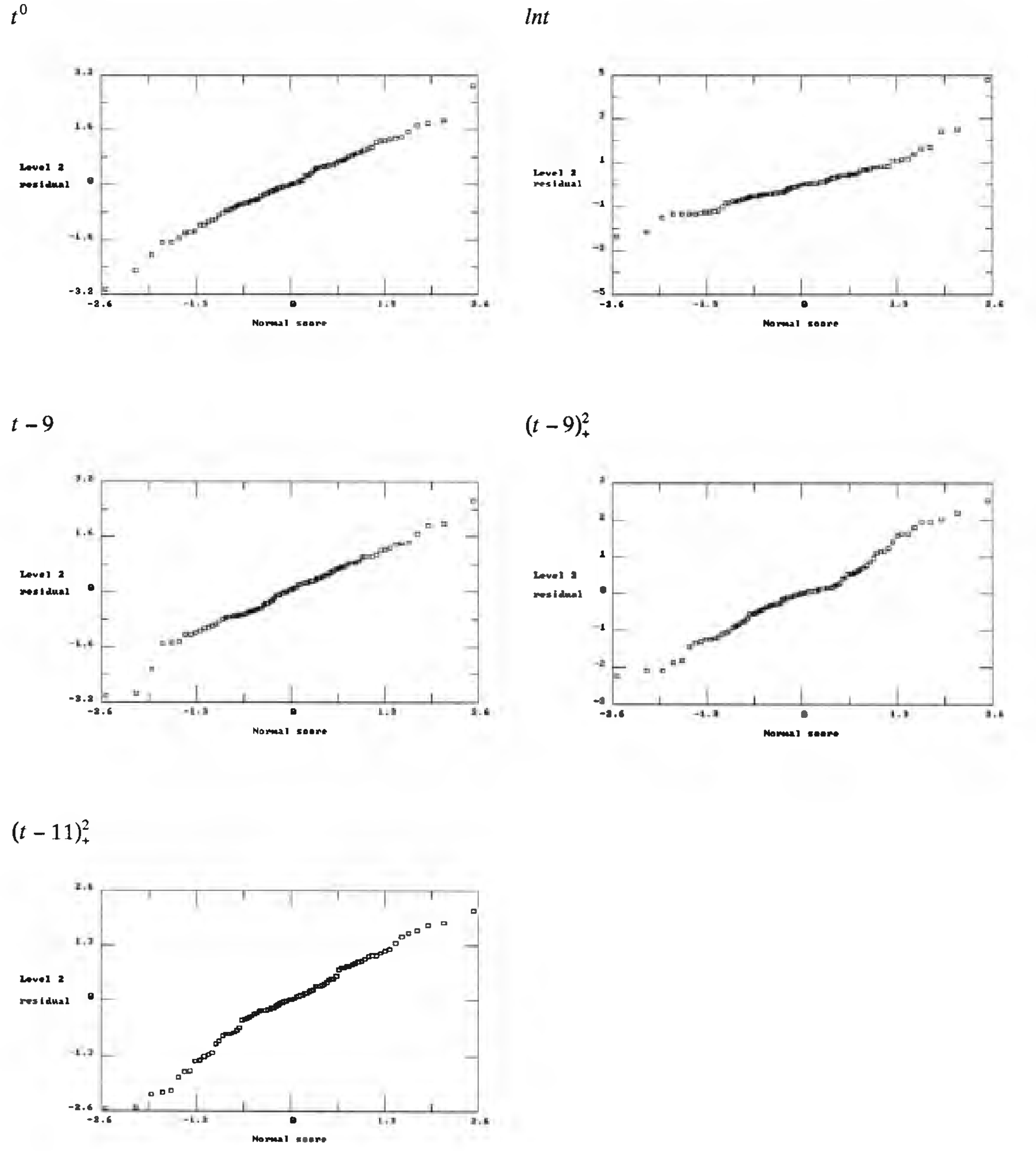


Figure 4.2.4 The Standardised level 2 residuals by Normal equivalent scores for the model in Table 4.2.6.



t^0

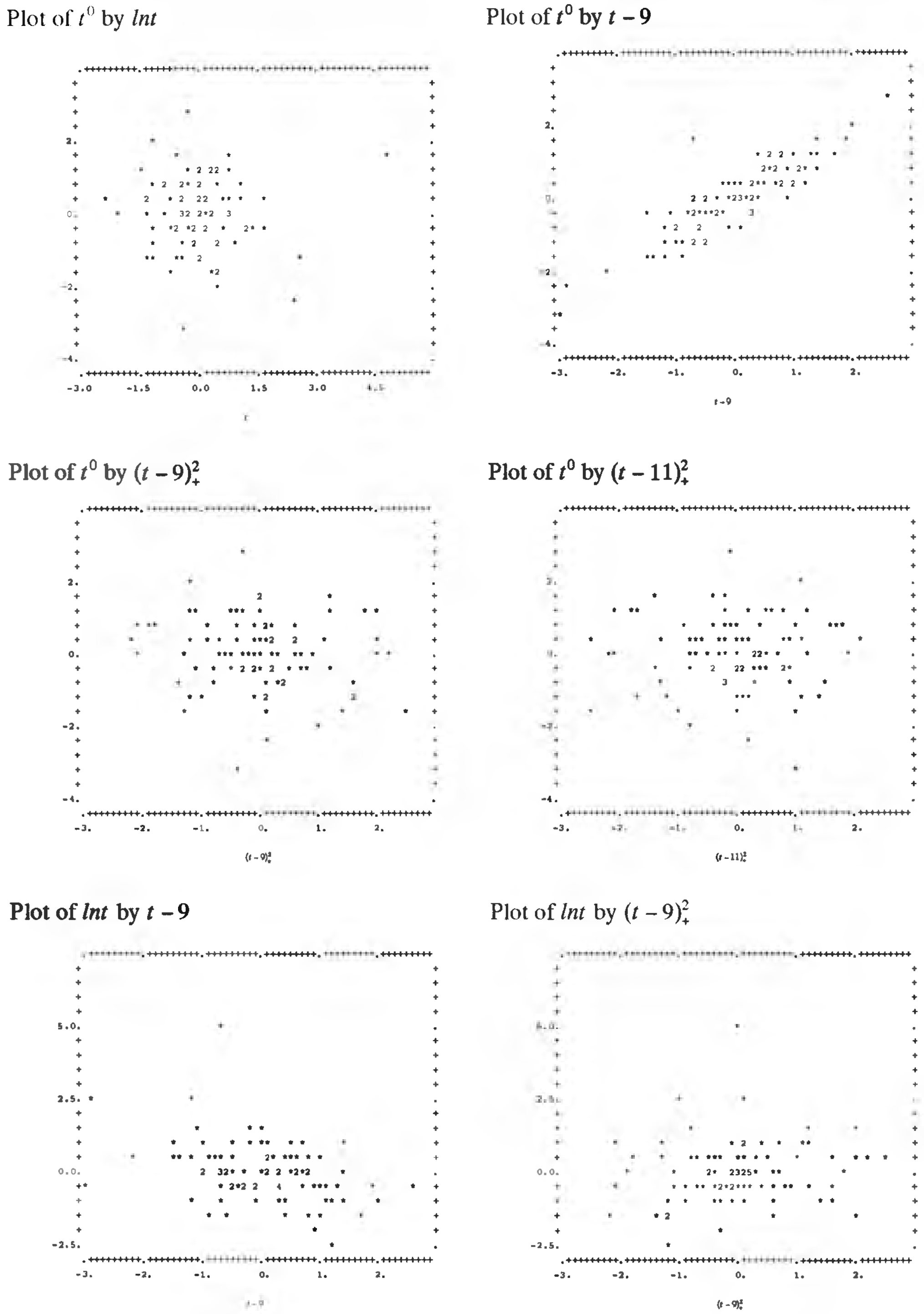
lnt

$t - 9$

$(t - 9)_+^2$

$(t - 11)_+^2$

Figure 4.2.5 Plots of standardised level 2 residuals.



Plot of t^0 by *lnt*

Plot of t^0 by $t-9$

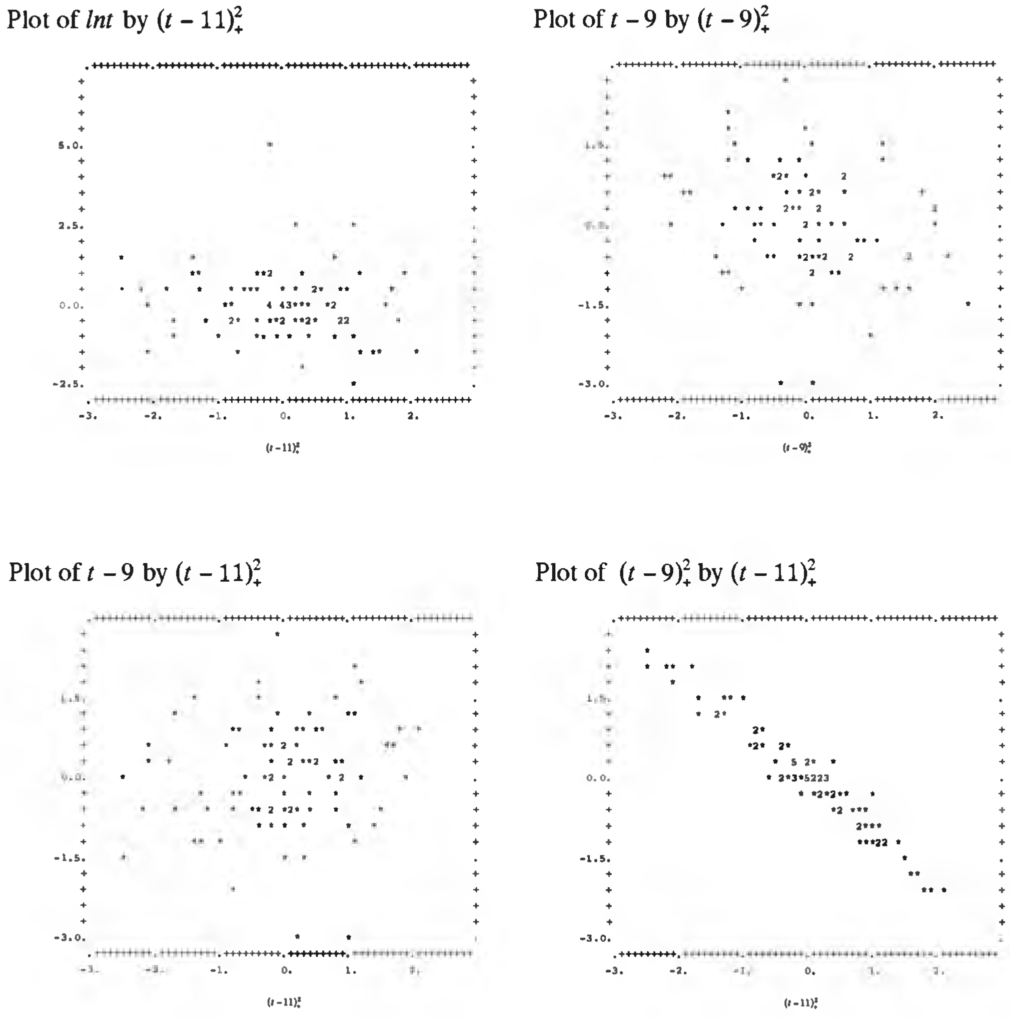
Plot of t^0 by $(t-9)^2_+$

Plot of t^0 by $(t-11)^2_+$

Plot of *lnt* by $t-9$

Plot of *lnt* by $(t-9)^2_+$

Figure 4.2.6 Plots of standardised level 2 residuals.



Plot of *lnt* by $(t-11)^2_+$

Plot of $t-9$ by $(t-9)^2_+$

Plot of $t-9$ by $(t-11)^2_+$

Plot of $(t-9)^2_+$ by $(t-11)^2_+$

The longitudinally estimated mean population curve of the model in Table 4.2.6 together with the cross-sectional means is shown in Figure 4.2.7. The two means are consistent in length. The cross-sectional means are derived from varying numbers of observations within the age intervals. The mean curve estimated by the multilevel model uses the precise age at which the measurement was taken. Figure 4.2.8 presents the estimated velocity curve which is the first derivative of the multilevel model based on the estimated fixed parameters.

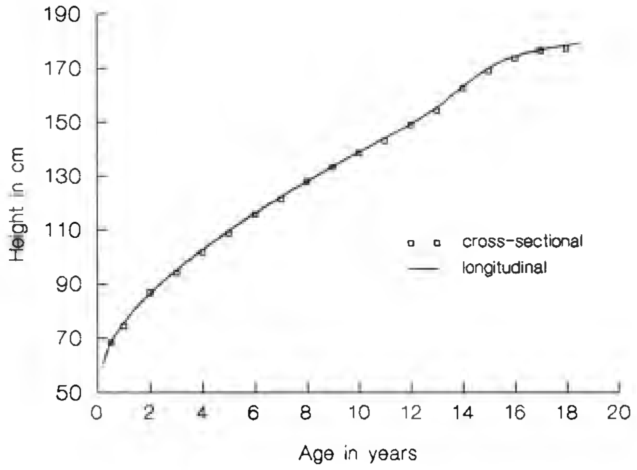


Figure 4.2.7 Longitudinally estimated mean curve (line) of HT for males with the cross-sectional means (dot).

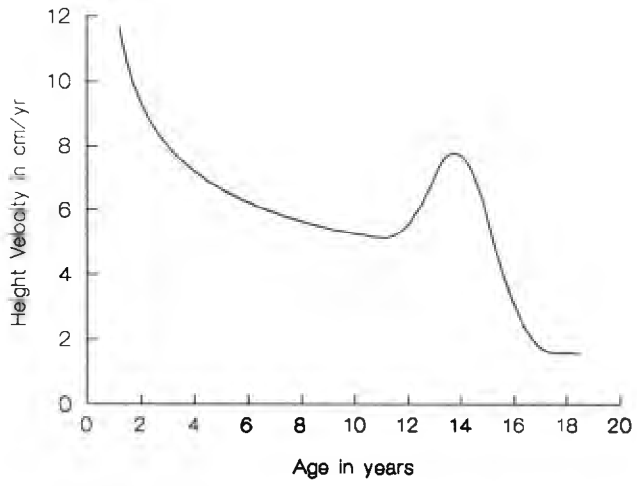


Figure 4.2.8 Estimated velocity curve of HT for males.

MODELLING HEIGHT FOR CONTROL FEMALES

The model (3.5) of Model HT-A in section 3 is investigated by fitting OLS curves for each individual. The knots at years 9, 11, 13, 15, 17 are chosen according to rules of thumb (Wold, 1974). The average residual standard deviation is 0.4904 cm with the range from 0.2347 cm to 0.7986 cm. Table 4.2.7 gives the correlation coefficients, means and standard deviations of the OLS estimates, where the standard error of g1 is equal to 0.2908 and of g2 is 0.5780.

The residual mean square error (RMS) ranges from 0.0551 cm^2 to 0.6377 cm^2 at average value of 0.2501 cm^2. For further check of the model, summary of residuals by age intervals is displayed in Table 4.2.8. The results are also close to expectation as were those for males.

Variance component models of height for 67 control females using polynomials, conventional splines of equation (2.18) and extended splines of equation (3.4) are illustrated in Figure 4.2.9 in HT-1(*), HT-2(*) and HT-3(*) respectively.

Comparing the values of LH and σ_e^2 for the models with seven or eight parameters in HT-1(*), HT-2(*) and HT-3(*), we can see that the extended splines fit the data better than the conventional splines. Comparing HT-3(7) with HT-3(8) we can also see that the term $(15-t)_+^3$ in HT-3(8) is not significantly different from zero ($p > 0.05$), that is, the knot at age 15 is not necessary in the population mean curve for females.

Table 4.2.9 shows values of LH and σ_e^2 of models derived from HT-3(7). The models with last knot at 17.5 years, such as models 1 and 2, do not show an improvement over the model of HT-3(7). The term $(15-t)_+^3$ is not significantly different from zero when comparing model 5 with HT-3(7). Models 6 to 13 are the model HT-3(7) with the first knot varying from age 1 to 8 years and they are close to model HT-3(7).

Table 4.2.7 Correlation, means and standard deviations of the OLS estimates

	t^0	lnt	$t-9$	$(9-t)_+^3$	$(11-t)_+^3$	$(13-t)_+^3$	$(15-t)_+^3$	$(17-t)_+^3$
Mean	149.900	4.8599	0.1307	0.0621	-0.3203	0.3124	0.0355	-0.1034
S.E.	14.6920	4.2186	1.3331	0.1617	0.3927	0.6240	0.5244	0.1666
g1	1.5377	-3.9926	1.3225	-0.3390	0.1838	-0.4345	0.1604	0.1012
g2	6.9145	21.2136	5.2984	-0.0979	-0.5698	-0.8313	-0.8798	-0.4461
Correlations								
t^0	1.0000							
lnt	-0.7177	1.0000						
$t-9$	-0.3131	-0.3171	1.0000					
$(9-t)_+^3$	-0.0927	0.3809	-0.3298	1.0000				
$(11-t)_+^3$	0.1334	-0.3038	0.1561	-0.7822	1.0000			
$(13-t)_+^3$	-0.2453	0.2082	0.1213	0.3194	-0.8296	1.0000		
$(15-t)_+^3$	0.3356	-0.1421	-0.3298	-0.0249	0.6048	-0.9425	1.0000	
$(17-t)_+^3$	-0.4101	0.1084	0.4768	-0.1266	-0.4458	0.8535	-0.9770	1.0000

Number of individual = 67 (females)

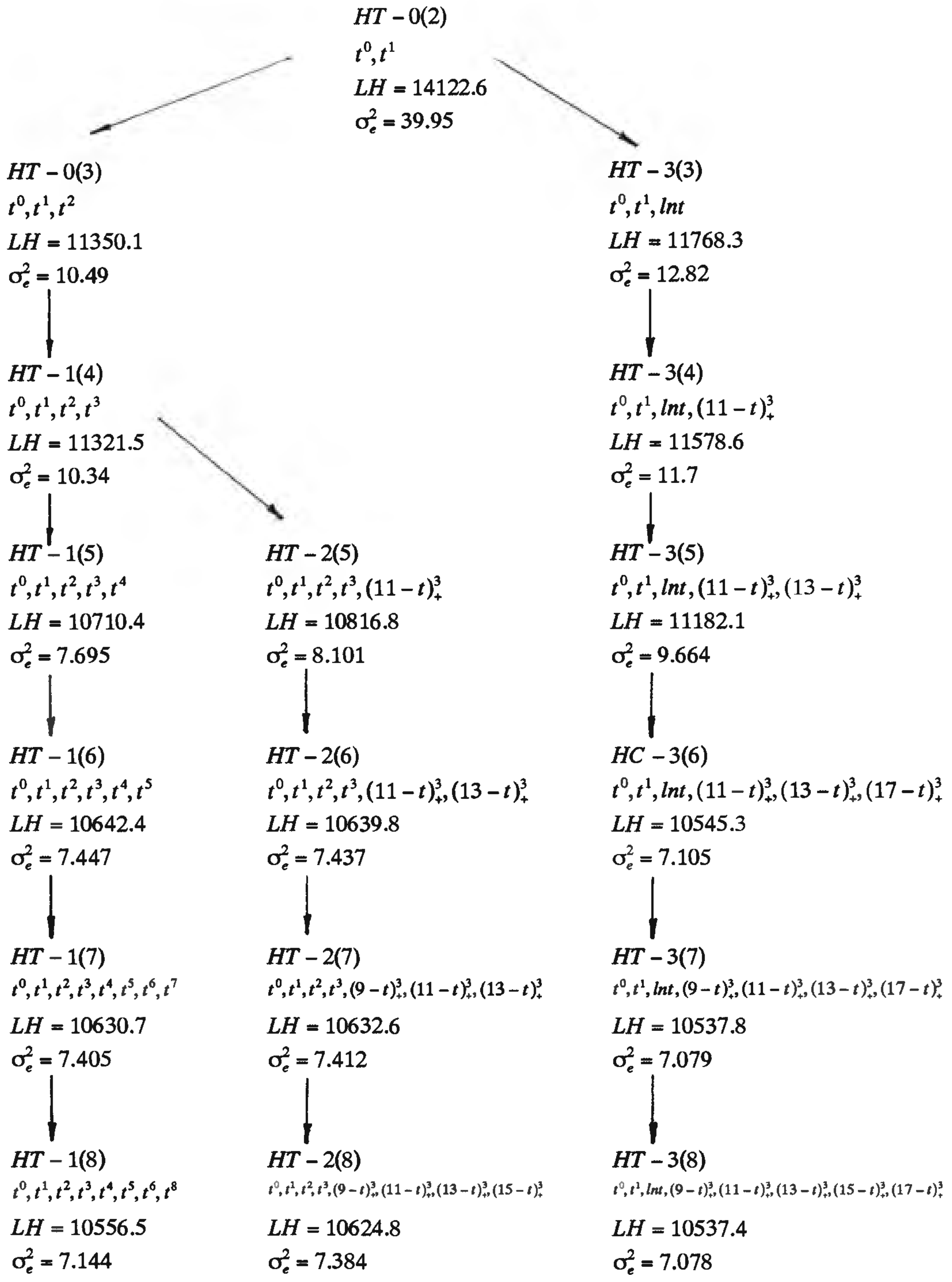
Note t^0 denotes intercept.

Table 4.2.8 Summary of HT residuals (OLS) by age group for females

Age	n	mean	sd	se	g1	g2
0.25+	62	-0.03	0.35	0.05	-0.12	0.58
0.50+	62	0.12	0.64	0.08	0.16	0.14
0.75+	56	0.01	0.60	0.08	-0.23	-0.33
1.00+	61	-0.03	0.65	0.08	0.35	1.84**
1.50+	59	-0.10	0.72	0.09	0.08	-0.58
2.00+	60	-0.10	0.57	0.07	-0.10	-0.45
2.50+	54	0.05	0.48	0.07	0.03	-0.57
3.00+	63	0.13*	0.48	0.06	-0.17	-0.87
3.50+	63	0.08	0.51	0.06	0.03	-0.24
4.00+	61	0.00	0.45	0.06	-0.15	-0.36
4.50+	67	-0.13**	0.44	0.05	-0.23	0.63
5.00+	59	0.07	0.45	0.06	-0.43	-0.07
5.50+	58	-0.02	0.56	0.07	-0.02	-0.46
6.00+	66	-0.10*	0.42	0.05	-0.36	0.78
6.50+	54	-0.06	0.43	0.06	-0.16	-0.35
7.00+	65	0.06	0.49	0.06	0.37	0.51
7.50+	57	0.00	0.51	0.07	-0.04	-0.12
8.00+	64	0.13**	0.40	0.05	-0.10	-0.27
8.50+	57	-0.00	0.38	0.05	-0.54	0.17
9.00+	68	-0.05	0.45	0.05	-0.03	-0.64
9.50+	57	-0.00	0.57	0.08	-0.94**	0.28
10.00+	62	-0.01	0.50	0.06	-0.73*	0.23
10.50+	60	-0.02	0.46	0.06	-0.37	0.80
11.00+	62	0.03	0.51	0.07	0.01	0.26
11.50+	65	0.03	0.61	0.08	0.64*	0.07
12.00+	59	-0.02	0.63	0.08	0.07	-0.83
12.50+	53	-0.02	0.44	0.06	-0.66*	0.25
13.00+	59	-0.01	0.51	0.07	0.61*	0.33
13.50+	61	-0.01	0.51	0.06	-0.28	-0.73
14.00+	60	0.04	0.47	0.06	0.10	-0.40
14.50+	54	0.00	0.42	0.06	-0.66*	1.86**
15.00+	60	-0.01	0.37	0.05	-0.46	1.05
15.50+	39	-0.05	0.42	0.07	-0.45	1.77*
16.00+	64	-0.02	0.41	0.05	-0.14	0.48
16.50+	40	0.06	0.44	0.07	1.31**	3.56**
17.00+	35	0.08	0.35	0.06	0.70	1.86*
17.50+	21	0.00	0.29	0.06	-0.77	1.07
18.00+	12	-0.16	0.59	0.17	-1.65**	3.54**

* p 0.05; ** P < 0.01

Figure 4.2.9 Variance component model of HT for 67 females



The model (3.6) of Model HT-B in section 3 is tested by including the term of $1/t$ into the model HT-3(7). The term $1/t$ is significant ($p < 0.05$) in model 2 comparing with model 1 in Table 4.2.9, however it does not improve much compared with model 3. Comparing model 3 with model 4 in Table 4.2.9 we can find that $1/t$ is not significantly different from zero ($P > 0.05$).

The random coefficient model of HC-3(7) is now used to analyse the data for females:

$$y_{ij} = \beta_{0j} + \beta_{1j}t_{ij} + \beta_{2j}\ln(t_{ij}) + \beta_{3j}(9 - t_{ij})_+^3 + \beta_{4j}(11 - t_{ij})_+^3 + \beta_{5j}(13 - t_{ij})_+^3 + \beta_{6j}(17 - t_{ij})_+^3 + e_{ij}, \quad (4.3)$$

$$\beta_{0j} = \gamma_0 + u_{0j} + u_{3j}(t_{ij} - 11)_+^2 + u_{4j}(t_{ij} - 13)_+^2,$$

$$\beta_{1j} = \gamma_1 + u_{1j},$$

$$\beta_{2j} = \gamma_2 + u_{2j},$$

$$\beta_{3j} = \gamma_3,$$

$$\beta_{4j} = \gamma_4,$$

$$\beta_{5j} = \gamma_5,$$

$$\beta_{6j} = \gamma_6.$$

Table 4.2.10 gives the estimates of parameters for the (4.3). Examining fixed coefficient contrasts shows that all the fixed coefficients are significantly different from zero ($P < 0.01$) except the coefficient of age.

Table 4.2.9 Model HT-3(7) with different knots or terms for females

	Models	LH	σ_e^2
1	$t^0, t^1, lnt, (9-t)_+^3, (11-t)_+^3, (13-t)_+^3, (17.5-t)_+^3$	10540.4	7.088
2	$t^0, t^1, lnt, (9-t)_+^3, (11-t)_+^3, (13-t)_+^3, (17.5-t)_+^3, 1/t$	10536.3	7.074
3	$t^0, t^1, lnt, (9-t)_+^3, (11-t)_+^3, (13-t)_+^3, (17-t)_+^3$	10537.8	7.079
4	$t^0, t^1, lnt, (9-t)_+^3, (11-t)_+^3, (13-t)_+^3, (17-t)_+^3, 1/t$	10534.1	7.066
5	$t^0, t^1, lnt, (9-t)_+^3, (11-t)_+^3, (13-t)_+^3, (15-t)_+^3, (17-t)_+^3$	10537.4	7.078
6	$t^0, t^1, lnt, (8-t)_+^3, (11-t)_+^3, (13-t)_+^3, (17-t)_+^3$	10537.3	7.078
7	$t^0, t^1, lnt, (7-t)_+^3, (11-t)_+^3, (13-t)_+^3, (17-t)_+^3$	10537.5	7.078
8	$t^0, t^1, lnt, (6-t)_+^3, (11-t)_+^3, (13-t)_+^3, (17-t)_+^3$	10537.9	7.080
9	$t^0, t^1, lnt, (5-t)_+^3, (11-t)_+^3, (13-t)_+^3, (17-t)_+^3$	10538.4	7.081
10	$t^0, t^1, lnt, (4-t)_+^3, (11-t)_+^3, (13-t)_+^3, (17-t)_+^3$	10538.7	7.082
11	$t^0, t^1, lnt, (3-t)_+^3, (11-t)_+^3, (13-t)_+^3, (17-t)_+^3$	10539.5	7.085
12	$t^0, t^1, lnt, (2-t)_+^3, (11-t)_+^3, (13-t)_+^3, (17-t)_+^3$	10545.1	7.104
13	$t^0, t^1, lnt, (1-t)_+^3, (11-t)_+^3, (13-t)_+^3, (17-t)_+^3$	10538.3	7.081

Table 4.2.10 Model (4.3) of height for 67 females

Fixed coefficient	Estimate	s.e.
t^0	146.2800	1.7650
lnt	5.7125	0.2875
$t-9$	0.2608	0.1723
$(9-t)_+^3$	0.0536	0.0125
$(11-t)_+^3$	-0.3034	0.0226
$(13-t)_+^3$	0.3233	0.0151
$(17-t)_+^3$	-0.0861	0.0033

Random parameters

Level 2 Covariance matrix (correlations in brackets)

	t^0	lnt	$t-9$	$(t-11)_+^2$	$(t-13)_+^2$
t^0	42.8170				
	(1.00)				
lnt	-4.8029	2.4207			
	(-0.47)	(1.00)			
$t-9$	4.5174	-0.7392	0.5360		
	(0.94)	(-0.65)	(1.00)		
$(t-11)_+^2$	-2.5608	0.3325	-0.3077	0.3193	
	(-0.69)	(0.38)	(-0.74)	(1.00)	
$(t-13)_+^2$	2.5792	-0.2154	0.3162	-0.4461	0.6990
	(0.47)	(-0.17)	(0.52)	(-0.94)	(1.00)
S.E. of Var.	7.6080	0.4581	0.0955	0.0578	0.1315

Level 1 variance = 0.9374 (0.0311)

Number of subjects = 67

Number of measurements = 2134

The likelihood ratio test is significant (p < 0.001) when comapring the model in Table 4.2.10 with its corresponding variance component model. The estimated standardised level 1 residuals' skewness, $g_1 = -0.0711$ (0.0530), and kurtosis, $g_2 = 1.1288$ (0.1059), show symmetry and a sharp peakedness of the distribution. For a check of the model, the plot of standardised residuals by predicted values is given in Figure 4.2.10, in which the residuals are randomly distributed and the residual variance increases with the magnitude of the predicted values, as for males, but no function could be found to describe this. The Normal plot of standardized level 1 residuals is displayed in Figure 4.2.11, showing an approximately Normal distribution.

Similarly to the male cases, the terms of '+' function are used to explain between-individual variation for the model in Table 4.2.10. The terms of $(t - 11)_+^2$ and $(t - 13)_+^2$ are finally chosen to form level 2 random part together with other variables, t^0, *lnt* and $t - 9$. For a check of level 2 residuals, the Normal plots of the standardised level 2 residuals by Normal equivalent scores are displayed in Figure 4.2.12. The plots of each pair of standardised level 2 residuals are shown in Figures 4.2.13-4.2.14.

Figure 4.2.10
Plot of standardised level 1 residual
by predicted values

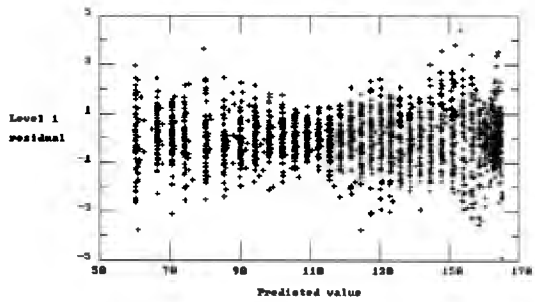


Figure 4.2.11
Plot of standardised level 1 residuals by
Normal equivalent scores

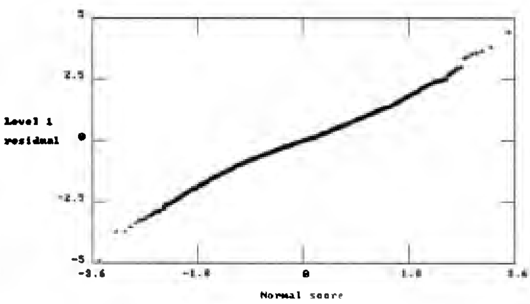
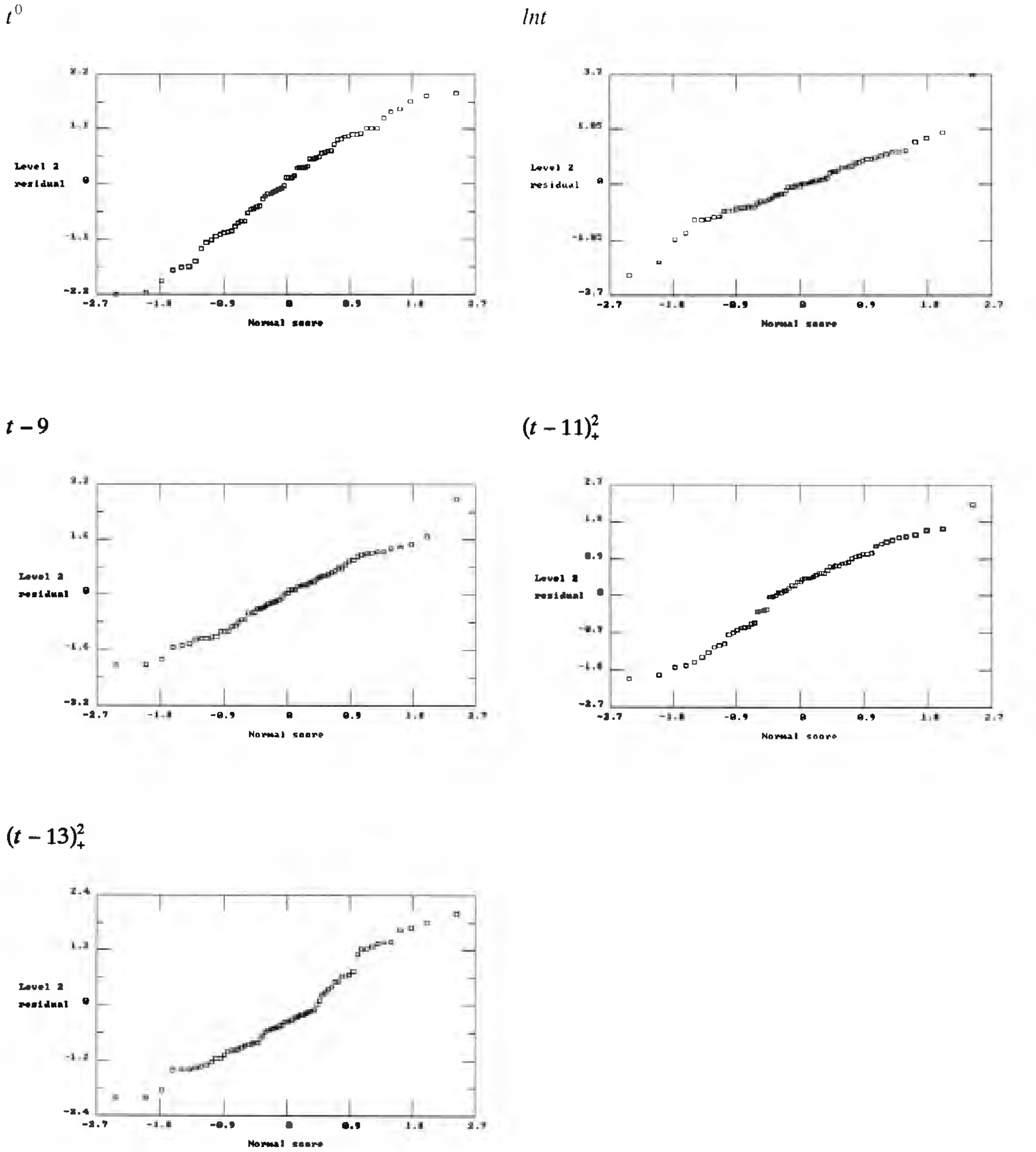


Figure 4.2.12 The Standardised level 2 residuals by Normal equivalent scores for Table 4.2.10.



t^0

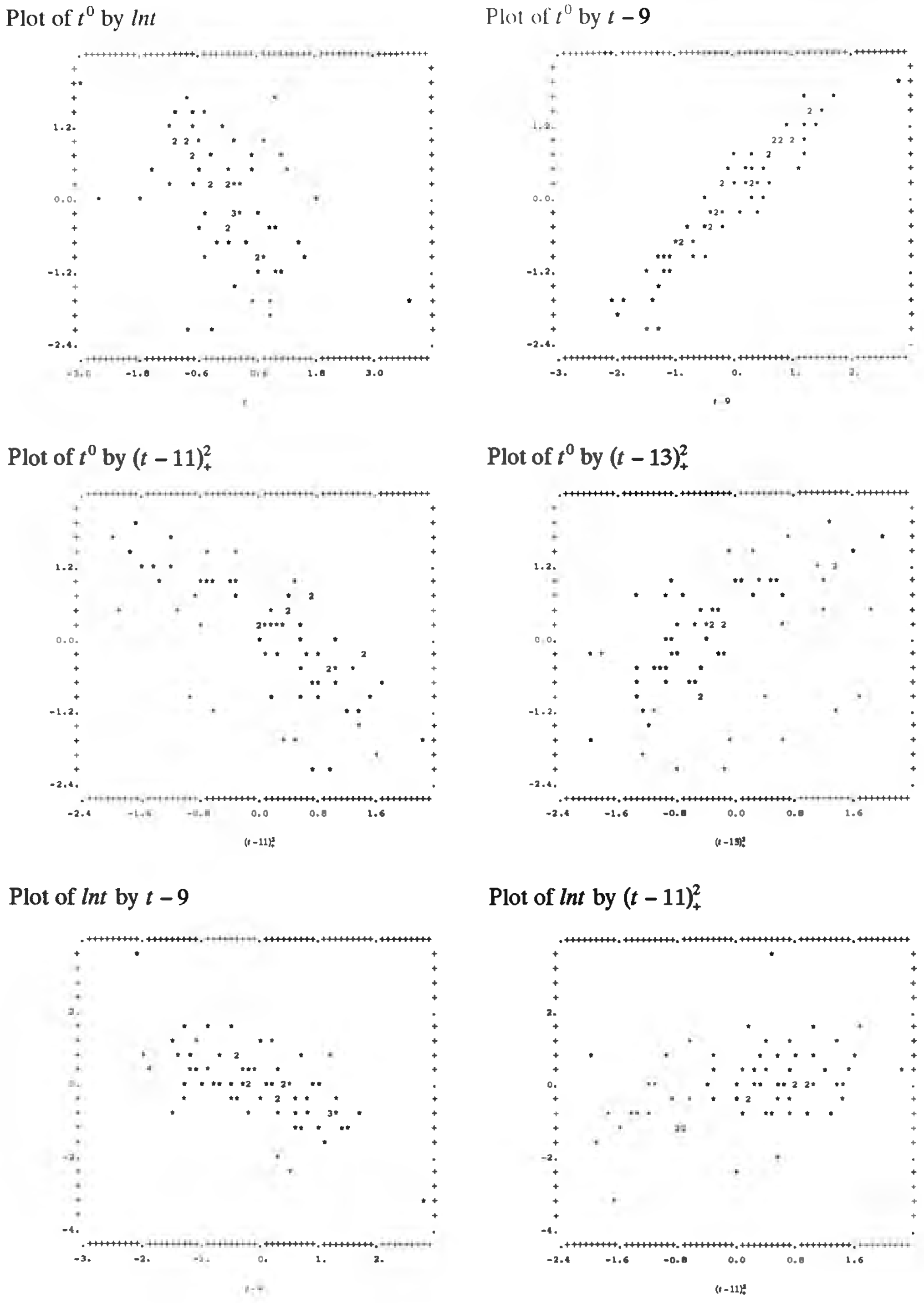
lnt

$t-9$

$(t-11)_+^2$

$(t-13)_+^2$

Figure 4.2.13 Plots of standardised level 2 residuals.



Plot of t^0 by *lnt*

Plot of t^0 by $t - 9$

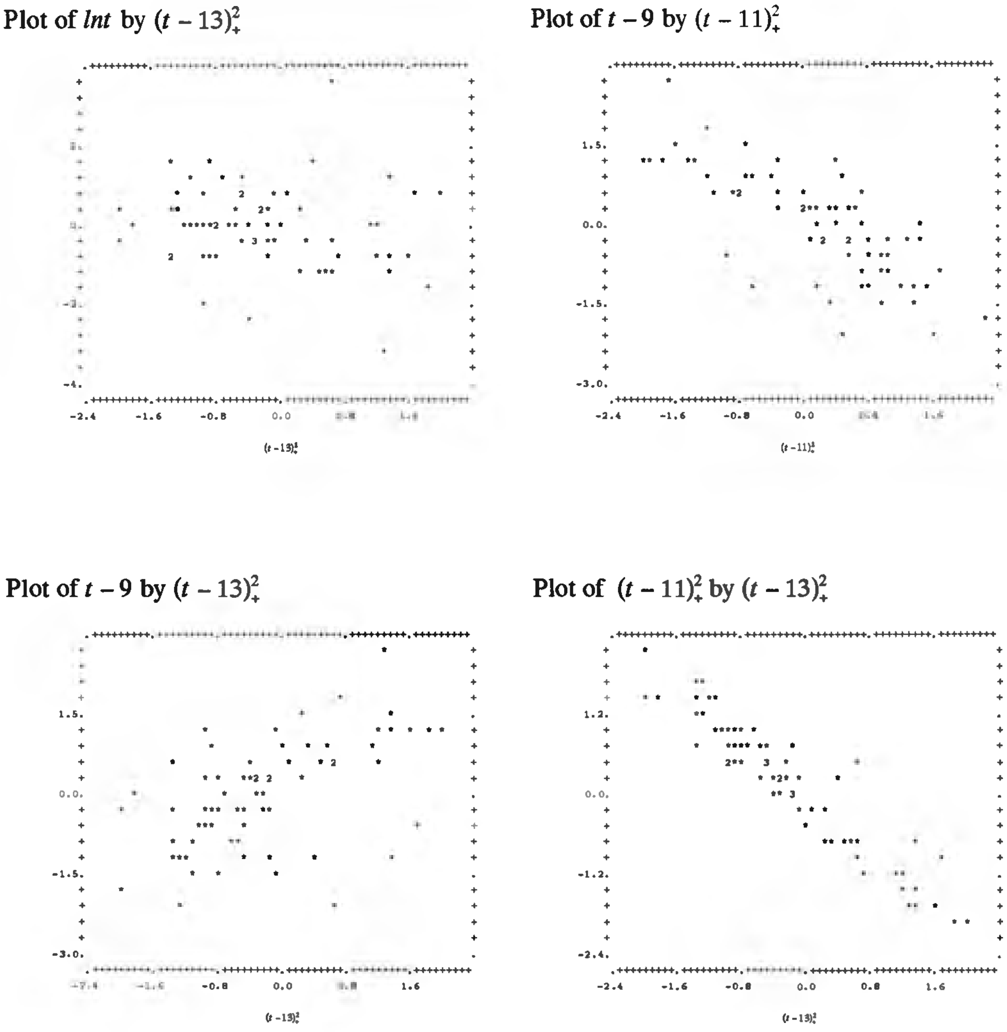
Plot of t^0 by $(t - 11)_+^2$

Plot of t^0 by $(t - 13)_+^2$

Plot of *lnt* by $t - 9$

Plot of *lnt* by $(t - 11)_+^2$

Figure 4.2.14 Plots of standardised level 2 residuals.



Plot of *lnt* by $(t-13)_+^2$

Plot of $t-9$ by $(t-11)_+^2$

Plot of $t-9$ by $(t-13)_+^2$

Plot of $(t-11)_+^2$ by $(t-13)_+^2$

Figure 4.2.15 shows the longitudinally estimated mean population curve of the random coefficient model of Table 4.2.10 together with the cross-sectional means. Figure 4.2.16 presents the estimated velocity curves from the model using the first derivative of the model based on the estimated fixed parameters.

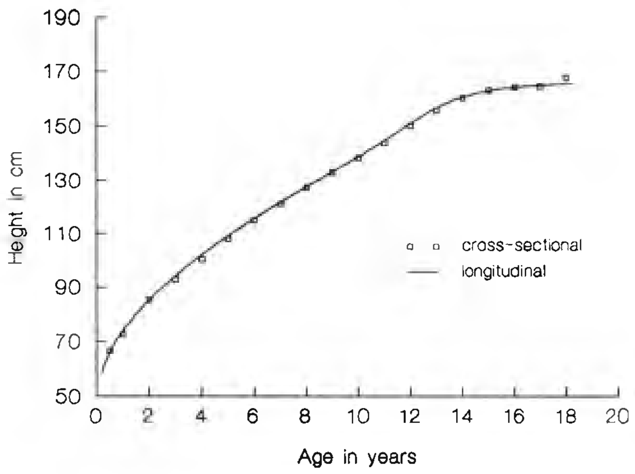


Figure 4.2.15 Longitudinally estimated mean curve (line) of HT for females with the cross-sectional means (dot).

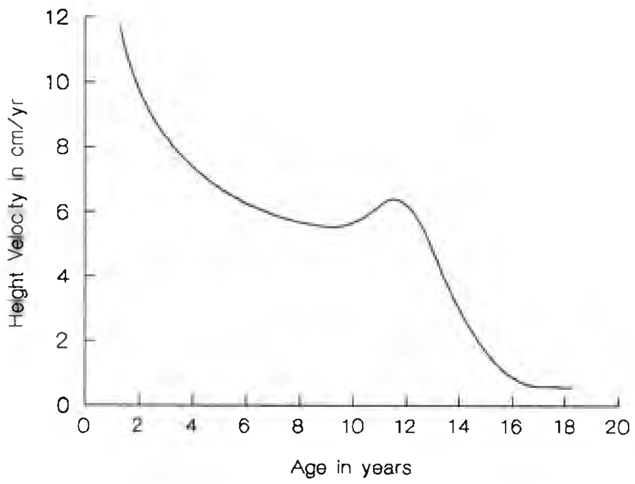


Figure 4.2.16 Estimated velocity curve of HT for females.

Growth Parameters

We have used the estimated parameters to calculate the velocity curves for males and females (see Figure 4.2.8 and 4.2.16). We can also use the estimated parameters to calculate growth parameters, such as take-off - at the minimum pre-spurt velocity, Peak Height Velocity (PHV) - maximal peak height velocity. These growth parameters are often compared between populations (Tanner, Whitehouse, Marubini and Resele, 1976).

If we use the model (4.2) for females as an example, then the ages of minimum or maximum velocity are given by the solution to the following equation:

$$-(\gamma_2 + u_{2j})/t^2 + 6\gamma_3(9 - t)_+ + 6\gamma_4(11 - t)_+ + 6\gamma_5(13 - t)_+ + 6\gamma_6(17 - t)_+ + 2u_{3j} + 2u_{4j} = 0$$

If we assume that the $-u_{2j}$, $2u_{3j}$ and $2u_{3j}$ have a multivariate Normal distribution we can estimate the distribution of t, for example, by simulation (Goldstein, 1989). In order to estimate the growth parameters properly, especially the variance of these parameters, we require at least coefficients up to the cubic to be random between individuals. However, it has not been able to include the random coefficients higher than the quadratic with the relatively small sample size. In addition, there is no closed form solution for the ages at minimum and maximum in the above equation. The values of ages at take-off and peak height velocity can be approximately read off from the velocity curves (Tanner, Whitehouse, Marubini and Resele, 1976; Berkey, Reed and Valadian, 1983).

Table 4.2.11 gives the mean values of the growth parameters. The means under the column 'Overall' are obtained using the estimates from the fixed part of Table 4.2.6 for males and Table 4.2.10 for females. Those under the column 'Simulation' are from a simulation sample of 100 individuals using the estimates, both the fixed and the random part, of these two tables. The two results are similar.

Table 4.2.11 Average growth parameters for 89 males and 67 females

Growth	Overall		Simulation	
Parameters	Males	Females	Males	Frmales
Age at take-off (years)	11.1	9.2	11.2	9.3
Velocity at take-off (cm/year)	5.1	5.5	4.9	5.6
Height at take-off (cm)	144.4	134.1	145.5	135.5
Age at PHV (years)	13.8	11.5	13.8	11.6
PHV (cm/year)	7.8	6.4	7.8	6.7
Height at PHV (cm)	161.4	147.6	161.6	149.2
Height at 18 years (cm)	178.1	165.1	178.9	165.5

Our results are very close to those obtained by other methods. For males, these estimates are close to the results of Ratcliffe, Pan and McKie (1992) for 16 boys of this dataset by using the kernel estimation (Gasser, Köhler, Müller, Kneip, Largo, Molinari and Prader, 1984; Gasser, Kneip, Ziegler, Largo and Prader, 1990). For females, the estimates are close to the results of Ratcliffe, Pan and McKie (1994) by using the same kernel estimation for 16 girls of this dataset. See Table 4.2.12 for the results using kernel estimation.

Table 4.2.12 Average growth parameters for 16 males and 16 females

Parameters	Males	Females
Age at take-off (years)	11.3 (1.0)	9.2 (1.6)
Velocity at take-off (cm/year)	4.7 (0.7)	5.0 (0.8)
Height at take-off (cm)	143.8(8.0)	134.1(10.3)
Age at PHV (years)	13.8 (1.2)	11.8 (1.1)
PHV (cm/year)	8.9 (1.3)	7.8 (0.8)
Height at PHV (cm)	159.6(7.6)	150.1(6.4)
Height at 18 years (cm)	175.1(7.7)	164.5(6.1)

It is worth noticing that the estimates of velocities at PHV by using the kernel estimation of Gasser, Kneip, Ziegler, Largo and Prader (1990) are higher than those obtained by any other parametric method so far. Their approach consists of synchronizing the individual curves before determining the average, which is biologically similar to the procedure of Tanner, Whitehouse and Takaishi (1966).

Our estimates for the age PHV are very close to those quoted by Bock and Thissen (1976) (see Table 4.2.13).

Table 4.2.13 Ages at adolescent PHV from various studies

Reference	Males	Females
Deming (1957)	13.4	11.4
Marubini et al (1972)	14.0	11.7
	14.2	11.9
Tanner et al (1976)	14.4	11.5
Bock et al (1973)	13.0	11.0
Bock et al (1976)	13.9	11.7

If we compare our results with those of the Fels Longitudinal Growth Study fitted with the triple logistic model by using maximum marginal likelihood estimation (Bock, 1992): the third logistic component is located at 13.75 years for male and 11.51 years for female and the corresponding growth velocities are 7.36 cm/year and 6.51 cm/year, which do not differ greatly from our estimates for velocity at these ages.

It is obvious that our models are better than those proposed by Berkey, Laird, Valadian and Gardner (1989). The five parameter linear Reed model was used for 62 boys aged from 8 to 18 years. The residual standard deviations of ordinary least squares (OLS) curves ranged from 0.37 to 2.51 cm with median of 1.4cm comparing 0.55 cm of our extended spline. The population mean was estimated by using the general linear random-effects model of Laird and Ware (1982). The ages both at take-off and PHV are underestimated by 0.5 year and the PHV is about 7 cm/year in their population mean curves for males.

Up to now, the eight parameter Reed model of Reed and Berkey (1989) is the only published linear model for height from birth to maturity. The model was used to fit curves for two boys and two girls from birth to 18 years of age. The plotted velocity curve has an obvious sudden change in it's shape at the joint, which implies the discontinuity in acceleration. In contrast, the velocity curves of Figure 4.2.8 and 4.2.16 represent the main features of the pubertal spurt well and are smooth at joints because these extended splines have three continuous derivatives at joints.

The height mean curves in Figures 4.2.7 and 4.2.15 have an asymptotic end for female but not for males. There may be two reasons for this: one is that females stop growing earlier than males and by year 18.5 some boys in late development are still growing; another is that relatively small numbers in the end age group might lead to an average curve being unstable at that age as discussed by Cole and Green (1992). In order to check the predicted values for males of 178.9 cm at 18.5 years using the 2-level random coefficient model with the extended spline, we applied BTT model to the data using the AUXAL program of Bock, Toit and Thissen (1994). The predicted value is 178.8 cm, which is close to our results. The BTT model is a nonlinear and asymptotic model, an extention to the Triple Logistic model of Bock and Thissen (1976).

Between-individual variation

Figure 4.2.17 shows the estimated between-individual standard deviation by age for each gender of the models in Table 4.2.6 and 4.2.10. For most ages, from early childhood to 16 years, these are close to those found by Tanner, Whitehouse and Takaishi (1966). The standard deviation increases by age gradually at the prepubertal stage and increase obviously at the pubertal stage. The difference in timing between the two genders is shown in Figure 4.2.17 and is reasonably consistent with the timing difference of the pubertal stages in gender. Note that the standard deviation increases after age 16 in our results which is not consistent with those of Tanner, Whitehouse and Takaishi (1966). This may be due to the relatively small sample and the increasing number of missing values at this age. The missing values vary from 0-18% in males and 0-42% in females at ages 0.25 to 16 years and from 9-63% in males and 4.5-82% for females after ages of 16 years. We have reasonably good estimates for the fixed coefficients. However, the estimates of the random terms depend on adequate number of individuals (Goldstein, 1986a).

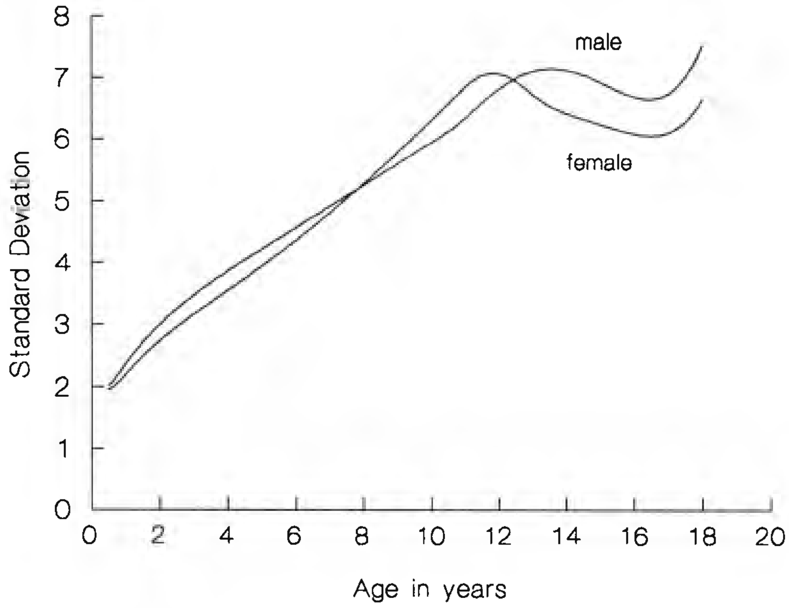


Figure 4.2.17 Between-individual standard deviation for Tables 4.2.6 of males and Table 4.2.10 of females

Chapter 5

Discussion

In longitudinal studies, growth patterns are often summarized by certain linear or nonlinear growth models so that a small number of parameters, or functions of them can be used to make group comparisons or to relate to other measurements. The statistical methods for analysis of longitudinal data can be described in two broad categories: fitting the average growth curve and fitting individual curves.

To analysis data based on the individual, growth models are fitted to each individual and then the growth parameters that describe the timing, magnitude and duration of the growth spurt are derived from the fitted models. The effects of covariates other than age, such as gender, protein etc. are often analysed by performing multiple cross-sectional analyses (Guo, Siervogel, Roche and Chumlea, 1992; Ratcliffe, Masera, Pan and McKie, 1994).

Multivariate analysis of variance with an unstructured variance-covariance matrix (Rao, 1965) uses polynomials and requires complete and balanced data. The application of multilevel models for describing growth is an efficient way for specifying growth models, linear or nonlinear, and incorporates covariates readily (Goldstein 1986a, 1989).

The interest of this study is to explore suitable specifications for multilevel models to analysis human growth and development over wide age ranges. The work has been focused on modelling human growth in height and head circumference. Measurements of height and head circumference are used in this study as they are of general interest and each includes a difficult range to be fitted: puberty in height measurements and infancy in head circumference with rapid acceleration or deceleration. The investigation of polynomials, conventional splines and the extended splines proposed in this study shows that the extended splines are better than polynomials and conventional splines for this purpose. The extended splines are useful, flexible, and simple to incorporate into multilevel models for population study.

5.1 Usefulness

These multilevel models are the only linear models proposed to date to estimate mean growth curves over a wide age range with covariates other than age. Prior to this study work carried out by other authors is limited to certain periods of age range for height. The model proposed by Berkey and Laird (1986) deals with height measurements for the Jenss curve for early childhood with gender and protein as covariates. The multilevel models using polynomials presented by Goldstein (1986a) and Goldstein (1989) show an efficient statistical modelling of longitudinal data over age 6 to 11 years with gender as covariate and covering ages 10 to 18 years respectively. The approaches of Berkey, Laird, Valadian and Gardner (1989) predict mean population parameters by the covariates of protein for a sample of 62 boys aged 8-18 years using the Reed model. The Jenss model is suitable for early childhood and the Reed model for the stage of puberty. No literature has addressed a population study for head circumference using multilevel models. Polynomials are flexible but they are neither able to fit adolescent growth in height nor the early childhood growth in head circumference as the curves are not polynomial like during these periods.

Multilevel models solve the inference problem we encounter when performing multiple cross-sectional analysis of the effects of covariates along the age scale. For example, the model for head circumference (HC) analyse the effect of the karyotype on magnitude along the age scale by using likelihood ratio tests on the parameters estimated for the various karyotype groups. For the same data set Ratcliffe, Masera, Pan and McKie (1994) studied the HC difference between normal and chromosomally abnormal groups by t-test with the Bonferroni inequality adjustment for multiple comparison in age groups. The two methods reach the same conclusion in analysing the effect of the karyotype. However multilevel models use data more efficiently than multiple cross-sectional comparisons especially when there is a small number of measurements within each year and each karyotype group and the data are unbalanced and incomplete. Multilevel models utilise the precise age when the measurement was taken and also all the information from each individual to estimate the covariate effects and mean curves.

These multilevel models can be compared with the model proposed by Berkey, Laird, Valadian and Gardner (1989), which uses the two-stage model of Laird and Ware (1982) incorporating the Reed model. The extended splines cannot be compared easily with the Reed model as the former has seven parameters and the later five parameters; the former is used for data covering early childhood to adult while the later covers adolescence to adult. However we can compare the two models for some important growth parameters derived from the two models with those quoted from published papers. The five-parameter linear Reed model was used for 62 boys aged from 8 to 18 years and their results in general were good. The residual standard deviations of ordinary least squares (OLS) ranged from 0.37 to 2.51 cm with a median of 1.4cm, while with our extended spline method it ranged from 0.29 to 0.91 cm with a mean of 0.55 cm. The age at take-off was estimated to be 9.7 years and the age at PHV about 13 years by Reed model, which were underestimated by about 1.0 and 0.5 year respectively in comparison with other studies (see Table 4.2.13). The PHV was estimated to be 7.0 cm/year in their population mean curves versus our PHV of 7.8 cm/yr. Our models seem more accurate than the model of Berkey, Laird, Valadian and Gardner (1989).

Recent work of Royston and Altman (1994) on fractional polynomials may provide a valuable contribution to modelling growth data. The question is whether fractional polynomials can be alternatives to the extended splines in this study. Royston and Altman (1994) provide a unified description and a degree of formalization for fractional polynomials.

In general fractional polynomials are the combination of logarithm terms and power terms with integer and non-integer values. The family of fractional polynomials are supposed to have considerable flexibility for various kinds of data. For $x > 0$ a fractional polynomial with degree k and powers $p_1 \leq \ldots \leq p_k$ is defined as

$$\Phi_k(x;\theta,p) = \sum_{i=0}^{k} \theta_i H_i(X),$$

where

$$H_i(x) = x^{(p_i)}, \qquad if \quad p_i \neq p_{i-1}$$

$$= H_{i-1}(x)\ln x, \qquad if \quad p_i = p_{i-1}.$$

The authors suggest that candidate values of p are all possible m-tuples selected with replacement from a fixed set of $\{-2,-1,-0.5,0,0.5,1,2,3,\ldots\max(3,k)\}$. With its definition of fractional polynomials the five-parameter Reed model can be expressed as $\Phi_4(t;-2,-1,0,1)$ and the cubic can be expressed as $\Phi_3(t;1,2,3)$.

A variety of fractional polynomials have been searched where the number of parameters varies from 7 to 8 to see whether we can find a fractional polynomial which can fit adolescence better than our extended splines. The strategy to enumerate all k-tuples from the set $\{-2,-1,-0.5,0,0.5,1,2,3,\ldots\max(3,k)\}$ is a heavy computational burden (Royston and Altman, 1994). The search was done by fixing some of the powers of set of $\{-4,-3,-2,-1,0,1,2,3\}$ and varying the others. Ordinary least squares (OLS) were used to fit individual curves for male cases in the Section 4. Standard deviations of the pooled residuals are used for comparison between the models. Table 5.1.1 shows part of the models for height with relatively small standard deviations of the pooled residuals. We find that the fractional polynomial $\Phi_5(0,0.5,1,2,3)$ is quite close to the extended splines in section 4 for head circumference. However we have not found any fractional polynomials which fit the height data better than the extended splines.

Table 5.1.1 Standard deviation of residuals (OLS) for height

Model	s
$\Phi_7(t;-4,-3,-2,-1,0,1,1)$	1.5517
$\Phi_7(t;-3,-2,-1,0,0,1,1)$	1.4636
$\Phi_7(t;-2,-1,-1,0,0,1,1)$	1.4661
$\Phi_7(t;-2,-2,-1,0,0,1,1)$	1.4476
$\Phi_7(t;-2,-1,0,0,1,1,2)$	1.3020
$\Phi_7(t;-1,0,0,1,1,2,2)$	1.1632
$\Phi_7(t;0,0,1,1,2,2,3)$	1.1134
$\Phi_7(t;0,1,1,2,2,3,3)$	1.0479
$\Phi_7(t;1,1,2,2,3,3,4)$	0.9016
$\Phi_7(t;0,1,2,3,3,4,4)$	0.8811
$\Phi_7(t;0,1,2,3,4,4,5)$	0.8496
$\Phi_7(t;0,1,2,3,4,5,6)$	0.7952
$\Phi_7(t;1,2,3,4,5,6,7)$	0.7046

The crude search may not be optimal. Some models may be prohibited by high correlations between power terms. For example, the correlation between variables $1/t$ and lnt/t is -0.9956; t and $lnt*t$ is -0.9999; $1/t$ and $1/t^2$ is 0.9946; lnt/t and $1/t^2$ -0.9999; $(lnt)^2/t$ and $(lnt)^2/t^2$ 0.9947. The high correlations between these terms make it difficult to find a fractional polynomial for height including childhood and adolescence where at least 7 parameters are required. Berkey and Reed (1987) suggest including $1/t^3$ and $1/t^4$ into the five-parameter Reed model to form a general Reed model. However Reed and Berkey (1989) find that in practice this is dubious and suggest using piecewise model.

The five-parameter Reed model includes logarithmic term and negative integer powers. Royston and Altman (1994) have applied fractional polynomials to many data sets and usually find a model that is an improved fit in comparison with the conventional polynomial.

Using the definition of fractional polynomials and the '+' function we can express extended splines with m segments in a general form as

$$f(t) = \sum_{i=0}^{k} \theta_i H_i(t) + \sum_{j=1}^{m-1} \phi_j (t - \xi)_+^r,$$

where

$$\begin{aligned} H_i(x) &= x^{(p_i)}, && \text{if } \; p_i \neq p_{i-1} \\ &= H_{i-1}(x)\ln x, && \text{if } \; p_i = p_{i-1}. \end{aligned}$$

Note the '+' function can be either $(t - \xi)_+^r$ or $(\zeta - t)_+^r$ or both if necessary. When $\max(p_i) \geq r$ or a negative power or logarithmic term is included the $f(t)$ has $r - 1$ continuous derivatives of t at the joints. For example the last segment in the extended splines for height is a fractional polynomial of $\Phi_2(t;0,1)$ and the last but one is $\Phi_3(t;0,1,3)$ and these two segments are combined at the last point with continuous at function, first and second derivatives. The extended splines combine segments of fractional polynomials and therefore they provide more useful tools than a single fractional polynomial.

5.2 Computational Simplicity

As the extended splines are piecewise fractional polynomials they have the advantages of both fractional polynomials and splines. With the '+' function in the extended splines the continuity constraints are implicit (Wold, 1974; Smith, 1979), which makes it possible to fit using standard methods.

The examples in section 4 illustrate the simplicity of fitting the extended splines with multilevel models. The use of the '+' function in extended splines allows the data to be fitted by any suitable method such as iterative generalized least squares (IGLS) of multilevel models straightforward without extra constraints.

Extended splines are not as computationally efficient as B-splines (Wold, 1974; Eubank, 1984) especially when a large number of knots are specified. The reason that B-splines are not used in this study is that they are not straightforward in interpretation and for analysing effects of covariates other than age. In addition it is not clear whether B-splines are suitable when splines contain terms of logarithmic, reciprocal etc. Smoothing splines are flexible tools for the estimation of a smooth curve but they are not available with a multilevel framework at the present time.

An alternative to the '+' function is using constraints to combine segments smoothly at joints (Cox, 1971, Seber and Wild, 1989; Reed and Berkey, 1989). Goldstein and Pan (1992) illustrate the use of constraints to smooth individual and centile curves respectively. Multilevel models allow us to join several segments with constraints both on fixed part and random parts (Goldstein 1987). This provides us with a broad way to fit segments for different age ranges with smooth constraints on joints.

In the early work of this study an attempt was made to combine polynomials with other functions, for example, a model composed of two segments, the first segment in cubic and the second $\frac{1}{t-c}+\frac{1}{(t-c)^2}$, where the value of c was given. Smooth constraints were put on the function, the first derivative and the level 1 and level 2 random part. Height measurements of 110 boys aged from 3 to 18 years were used from the mixed data of the Harpenden Longitudinal Study and International Children's Centre Study (Tanner, Whitehouse, Marubini and Pesele, 1976; Tanner, Goldstein and Whitehouse, 1970). It was difficult for the iterative procedure to converge when three segments were included or when covariates other than age are considered because of the need for smooth constraints at joints both for the fixed part and the random part. This leads to too many parameters to be estimated when a limited number of level 2 units are available.

Similar problems may occur when the model of Reed and Berkey (1989) is incorporated into a multilevel model. The Reed and Berkey model is composed of two fractional polynomials: the first one is $\Phi_4(t; -2, -1, 0, 1)$ for the pre pubertal stage and the second $\Phi_4(t; -3, -2, -1, 0)$ for the pubertal stage. In principle it can be estimated by multilevel models with constraints. We need at least 5 constraints: three for the fixed part to be continuous at function, first and second derivatives and two for level 1 and level 2 random variance to be equal at joints, which leads computational difficulties especially if covariates other than age are required to be estimated. This problem may explain why there are few applications of the model despite the fact that the Reed and Berkey model is the first and the only linear model published which describes human growth from birth to maturity. The work of Reed and Berkey (1989) was focused on fitting individual curves by OLS with constraints for continuous on the function and the first derivative. The velocity curve derived from the model has an obvious sudden change in it's shape at the joint, which implies a discontinuity in acceleration. In contrast, the velocity curves derived from our extended splines present the main features of the pubertal spurt well and are smooth at the joints because these extended splines are continuous up to the second derivative at the joints (see Figure 4.2.8 and 4.2.16).

5.3 Flexibility

Most growth models describe growth patterns for certain periods. For example, the Reed model (Berkey and Reed, 1987) and Jenss model (Jenss and Bayley, 1937) are suitable for the height of young children; the triple logistic (Bock and Thisse, 1976), BTT model (Bock, Toit and Thissen, 1994), JPPS model (Jolicoeur, Pontier and Sempe, 1988) and JPA2 model (Jolicoeur, Pontier and Abidi, 1992) are suitable for height from early ages to maturity. They may not be able to fit data for other age ranges. In addition a model for height may not suit other measurements, for instance, head circumference or leg length. By contrast we can add or drop some terms from an extended spline or use different joints to form an extended spline for other measurements or other age ranges.

An example is given now to demonstrate the flexibility of extended splines. The extended splines for height can be applied to model other measurements of length, such as sitting height and leg length. The measurements of sitting height (SH) and leg length (LL) are from the 83 normal males and 60 females, from the same sample used in section 4.1. Sitting height was measured and leg length was obtained by subtracting sitting height from the corresponding total standing. The data cover ages 2 to 18 years. For a preliminary check, ordinary least squares (OLS) curves are fitted independently to each individual's data. The model (3.5) is used with knots set at 11, 13, 15 and 17 years for males and 9, 11, 13 and 17 years for girls respectively. The check of the residuals gives similar results as we found in the fitting for height (see section 4.1). The residual standard deviations are summarized in Table 5.3.1.

Table 5.3.1 Residual standard deviation for SH and LL (OLS)

		SH		LL	
Gender	N	Mean	Range	Mean	Range
Male	83	0.48	0.23 - 0.72	0.48	0.29 - 0.81
Female	60	0.48	0.25 - 0.79	0.47	0.32 - 0.78

The flexibility of the extended splines allows us to study changes of relationships between different kinds of measurements by ages using a multivariate longitudinal model. As the numbers of individuals are greater for males than for females, we will use the data of the 83 males as an example to show how a multivariate longitudinal model for sitting height and leg length can be formulated. Level 1 is the variable, sitting height or leg length within occasion which is level 2 and subject is level 3.

The values of the mean and standard deviation of sitting height and leg length from the raw data are listed for each age group in Table 5.3.2. Table 5.3.3 gives the estimates of fixed parameters using the extended splines. All the parameters in the fixed part are significantly different from zero ($P < 0.05$ or $P < 0.01$) except the variable age in the model for leg length. The parameters of the mean population curves are significantly different between these two curves ($P < 0.001$).

Figure 5.3.1 shows the mean curves for sitting height and leg length (in line) and also shows the mean values estimated cross-sectionally. The main growth parameters derived from the estimated fixed part are listed in Table 5.3.4. The growth parameters are close to those of Tanner, Whitehouse, Murubini and Resele (1976) and Gasser, Kneip, Binding, Prader and Monlinari (1991) despite using different methods and data.

Table 5.3.5 gives the covariance matrix of random coefficients of the multivariate longitudinal model. The likelihood ratio for this model is 13362.3 and is 17976.4 for the variance component model. The difference is highly significant ($p < 0.001$) and indicates that the variables in the level 3 random part are necessary.

Using the estimated parameters from the multivariate model we can obtain correlations at any ages of the data. Tables 5.3.6 and 5.3.7 show estimated correlations at specified ages for sitting height and leg length respectively. Table 5.3.8 gives the correlation between sitting height and leg length changes by age. On average it is 0.51 at two years of age and increases to 0.68 at ten years of age and decreases to 0.43 at eighteen years of age.

In practice longitudinal data are not measured exactly at the target ages. The multivariate longitudinal model provides an efficient way to analyse data in this case. Goldstein (1986) gives an example of a multivariate model using polynomials for height and weight at ages about 8 to 9 years to predict correlations for target ages 8 and 9 years. The example given here for ages 2 to 18 years shows that the extended splines are useful in order to formulate a multivariate model when data do not behave like polynomials and when data cover wide age ranges.

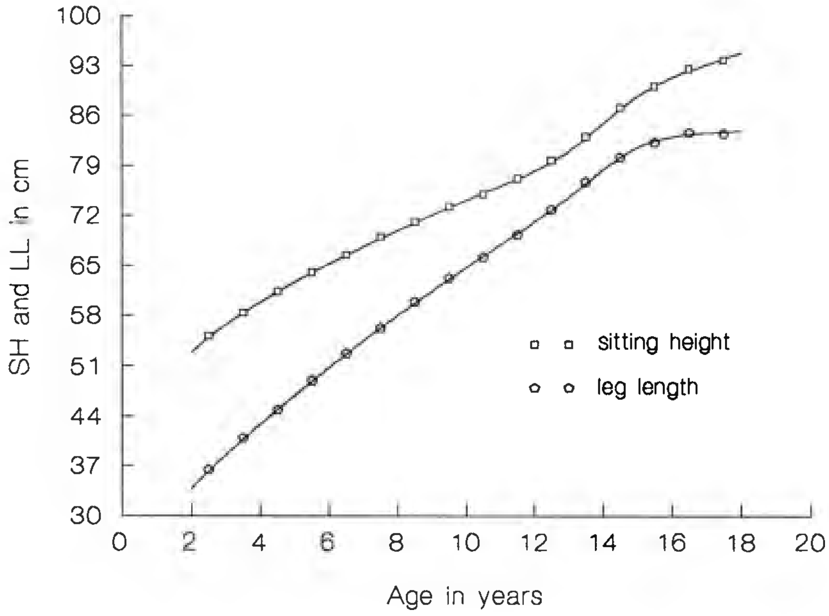


Figure 5.3.1 Mean curves of sitting height and leg length for Table 5.3.3 of males

Table 5.3.2 Means of SH and LL (cm) in age group for 83 males

		SH		LL	
Age Group	N	Mean	S.D.	Mean	S.D.
2.00+	67	53.24	1.94	34.12	1.91
2.50+	71	55.19	1.89	36.68	2.19
3.00+	83	56.95	2.06	38.95	2.17
3.50+	82	58.48	2.14	41.01	2.34
4.00+	86	60.37	2.07	43.20	2.53
4.50+	79	61.36	2.29	45.07	2.68
5.00+	84	62.99	2.37	47.43	2.75
5.50+	81	64.09	2.45	49.33	3.00
6.00+	73	65.42	2.38	51.06	3.01
6.50+	79	66.60	2.57	52.95	3.14
7.00+	80	67.61	2.41	54.61	3.13
7.50+	79	69.22	2.74	56.50	3.13
8.00+	77	70.08	2.60	58.25	3.46
8.50+	75	71.16	2.95	60.07	3.51
9.00+	85	72.23	2.60	61.85	3.73
9.50+	76	73.35	2.58	63.42	3.71
10.00+	81	74.20	2.88	65.14	3.77
10.50+	77	75.08	3.09	66.21	4.08
11.00+	78	76.20	2.68	68.01	3.86
11.50+	77	77.49	3.47	69.35	4.01
12.00+	76	78.40	3.28	71.52	4.16
12.50+	79	79.55	3.66	72.81	4.56
13.00+	74	81.57	4.34	74.85	4.77
13.50+	67	82.95	3.66	76.85	4.60
14.00+	80	85.44	4.31	78.89	4.65
14.50+	76	87.56	4.30	80.49	4.68
15.00+	73	88.39	3.80	81.41	4.42
15.50+	71	90.42	3.82	82.41	4.54
16.00+	77	91.67	3.49	82.89	4.69
16.50+	64	92.73	3.12	83.81	4.49
17.00+	56	93.12	3.13	83.47	4.35
17.50+	45	93.84	3.03	83.26	5.04

Table 5.3.3 Mean curves of sitting height and leg length for 83 males

Fixed coefficient	Estimate	s.e.
Sitting Height		
Intercept	66.7130	4.1670
$t-10$	1.3477	0.1572
lnt	5.9621	1.7290
$(11-t)_+^3$	0.1295	0.0089
$(13-t)_+^3$	-0.3840	0.0162
$(15-t)_+^3$	0.3840	0.0152
$(17-t)_+^3$	-0.1286	0.0058
Leg Length		
Intercept	74.1280	3.7850
$t-10$	0.1300	0.1582
lnt	2.9560	1.5620
$(11-t)_+^3$	0.0622	0.0081
$(13-t)_+^3$	-0.2893	0.0148
$(15-t)_+^3$	0.3952	0.0141
$(17-t)_+^3$	-0.1689	0.0055

Table 5.3.4 Main growth parameters of SH and LL for 83 males

Variable	Our results	Tanner et al (1976)	Gasser et al (1991)
Sitting Height			
Age at take off (yr)	11.20	12.12	11.30
Age at peak velocity (cm/yr)	14.00	14.25	14.20
Peak velocity (cm/yr)	4.09	4.54	4.84
Adult size (cm)	94.73	92.71	93.80
Leg Length			
Age at take off (yr)	11.30	12.01	10.70
Age at peak velocity (cm/yr)	13.50	13.58	13.80
Peak velocity (cm/yr)	3.89	4.25	4.87
Adult size (cm)	83.71	80.94	84.30

Table 5.3.5 Covariance Matrix of random coefficients (correlations in brackets)

		SH				LL			
		t^0	t	t^2	t^3	t^0	t	t^2	t^3
SH	t^0	8.9202 (1.00)							
	t	0.5945 (0.54)	0.1366 (1.00)						
	t^2	-0.0570 (-0.60)	-0.0032 (-0.27)	0.0010 (1.00)					
	t^3	-0.0078 (-0.44)	-0.0018 (-0.82)	0.0001 (0.54)	0.0000 (1.00)				
LL	t^0	8.2572 (0.72)	0.4504 (0.32)	-0.0413 (-0.34)	-0.0054 (-0.24)	14.6150 (1.00)			
	t	0.4528 (0.51)	0.0683 (0.62)	-0.0014 (-0.15)	-0.0007 (-0.39)	0.8030 (0.70)	0.0900 (1.00)		
	t^2	-0.0778 (-0.71)	-0.0091 (-0.67)	0.0007 (0.60)	0.0002 (0.92)	-0.0894 (-0.65)	-0.0058 (-0.53)	0.0013 (1.00)	
	t^3	-0.0074 (-0.43)	-0.0018 (-0.84)	0.0001 (0.55)	0.0000 (0.80)	-0.0070 (-0.32)	-0.0012 (-0.69)	0.0002 (0.74)	0.0000 (1.00)

Number of measurements = 2459 in sitting height.

Number of measurements = 1637 in leg length.

The within-subject variance is 0.6853 for sitting height.

The within-subject variance is 0.5653 for leg length.

t^0, t, t^2 and t^3 denote intercept, age-10, (age-10)^2 and (age-10)^3 respectively.

Table 5.3.6 Estimated correlations for sitting height

Age (year)	2	4	6	8	10	12	14	16
2								
4	0.75							
6	0.70	0.85						
8	0.71	0.82	0.87					
10	0.70	0.75	0.82	0.89				
12	0.68	0.67	0.74	0.85	0.92			
14	0.64	0.61	0.68	0.80	0.89	0.93		
16	0.55	0.56	0.62	0.71	0.80	0.85	0.91	
18	0.32	0.42	0.46	0.49	0.51	0.56	0.65	0.82

Table 5.3.7 Estimated correlations for leg length

Age (year)	2	4	6	8	10	12	14	16
2								
4	0.79							
6	0.77	0.91						
8	0.78	0.90	0.93					
10	0.79	0.88	0.92	0.95				
12	0.78	0.85	0.90	0.94	0.96			
14	0.75	0.84	0.88	0.92	0.95	0.96		
16	0.66	0.80	0.84	0.86	0.87	0.89	0.93	
18	0.40	0.63	0.65	0.63	0.61	0.62	0.71	0.87

Table 5.3.8 Estimated correlations for sitting height and leg length

					SH					
	Age (year)	2	4	6	8	10	12	14	16	18
	2	0.51	0.64	0.73	0.78	0.78	0.73	0.65	0.51	0.22
	4	0.44	0.64	0.72	0.74	0.71	0.66	0.62	0.57	0.45
	6	0.41	0.59	0.67	0.70	0.68	0.63	0.59	0.55	0.42
	8	0.42	0.56	0.64	0.68	0.68	0.64	0.59	0.51	0.31
LL	10	0.43	0.52	0.60	0.67	0.68	0.66	0.59	0.48	0.21
	12	0.44	0.49	0.57	0.65	0.68	0.66	0.60	0.46	0.15
	14	0.44	0.47	0.55	0.63	0.67	0.66	0.61	0.47	0.15
	16	0.42	0.44	0.50	0.58	0.62	0.62	0.59	0.50	0.25
	18	0.30	0.33	0.36	0.40	0.42	0.44	0.46	0.43	0.43

5.4 An Optimal Model

The procedure shown in chapter 4 does not guarantee an optimal model for the data studied. Five parameters are considered in describing a growth trajectory of head circumference and seven or eight parameters of height as other authors suggest (Bock and Thissen, 1980; Jolicoeur, Pontier and Abidi, 1992). There is a choice of fractional polynomials from the set of $\{-2,-1,0,1\}$ and a small number of knots. The placement of knots is determined according to the rule of thumb (Wold, 1974) on the basis of knowledge of growth or the experience of other authors. For example, if we count the number of maxima, minima and inflection points in a height curve we can see that we need a model with five knots, which are initially fixed at 9, 11, 13, 15 and 17 years and the other three terms chosen from $\{-2,-1,0,1\}$. We need to experiment on the placement of knots by moving one knot when the other four are fixed. Chapter 4 illustrates this. The final model will be determined by its interpretability and small residuals. With eight parameters the amount of computation for all possible models is very large. Experience indicates that there may exist several models which are close to the 'optimal' one and can describe the data well. Royston remarks (1994):

However, we hypothesize that the likelihood surface (with respect to p) will often be very flat for $m > 2$, making the choice of p less critical than with $m \leq 2$.

5.5 Individual and Population Curves

It should be noted that the mean growth curve does not necessarily have the same form as individual curves (Bryk and Raudenbush, 1987). In this study the five-parameter model is used for head circumference for both individual curves and mean curves in each gender. Eight-parameter models are used for individual height curves using OLS and seven-parameter models for mean curves using multilevel models. The '+' function of $(15-t)_+^3$ is not significant in the fixed part of the multilevel model for females and $(9-t)_+^3$ is not significant in the fixed part of the model for males. This is because of the greater variation in height both in timing and magnitude between individuals. If we plot mean growth curve overlapped by individuals curves we can see that some individual trajectories with positive curvatures cancel out others with negative curvatures. Multilevel models allow us to model this variation. For example, in the model for males we can let $(9-t)_+^3$ random at level 2 but have a small value (or zero) in the fixed part.

In the model for head circumference all the five variables in the fixed part are random at level 2 and the estimated within-individual variance is consistent with the variance of errors for individual curves of other studies (Roche, Mukherjee and Guo, 1986). In the model for height there are five random variables at level 2 and the estimated within-individual variance is around $1.0cm^2$. It is the default value of the variance of the measurement errors in the AUXAL Program (Bock, Toit and Thissen, 1994). It is slightly larger than the estimates of residual variance in our individual curves. We need a larger number of level 2 units in order to include more or higher-order random coefficients to model adequately between-individual variation in height to obtain smaller level 1 variance estimate. See the discussion of Goldstein (1986a).

5.6 Midgrowth Spurt

The evidence of the occurrence of a small growth spurt in height several years before the pubertal spurt has been cited by several authors (Stützle, Gasser, Molinari, Largo, Prader and Huber, 1980; Gasser, Müller, Köhler, Prader, Largo and Molinari, 1985, Goldstein, 1986a). These spurts have been described as the midgrowth spurt or mid-childhood growth spurt (MS). Bock and Thissen (1980) found a pronounced midgrowth spurt in the average height velocity curve for either gender of Berkley Guidance Study by differentiating the triple logistic average parameter curves. It should be noticed that a mean parameter curve of a nonlinear model is not the mean population curve (Merrell, 1937) therefore we should be cautious of the pronounced midspurt spurt in the mean parameter curve of the nonlinear triple logistic of Bock and Thissen (1980).

Gasser, Kneip, Ziegler, Largo and Prader (1990) showed the midgrowth spurt at about ages 6 to 9 years in boys and girls of Zurich Longitudinal Growth Study. They evolved a new statistical method to shift individual curves continuously (and non-linearly) in age prior to averaging the individual growth curves. Their resulting curves present the typical shape rather than individual variations in it.

The recent work of Bock, Toit and Thissen (1994) gives the estimation for nonstructure average using the Maximum A Posterior (MAP) procedure. Midgrowth spurts obtained are close to those shown by Gasser, Kneip, Ziegler, Largo and Prader (1990) and are not as pronounced as those given by Bock and Thissen (1980).

A number of authors have not observed the regular presence of the midgrowth spurt in measures of height velocity. Tanner and Cameron (1980) examined single-year increments from a sample of 10,000 children in London and did not find evidence for a midgrowth spurt in girls velocity curves and only demonstrated a diminution of deceleration from age six to seven years in boys. Meredith (1981) found a midgrowth spurt in 14% of Iowa City children.

Berkey and Reed and Valadian (1983) demonstrated a midgrowth spurt in height in 17 of 67 boys and 0 of 67 girls from Longitudinal Studies of Harvard University using the variable knot cubic splines. Jolicoeur, Pontier, Pernin and Sempe (1988) used a seven-parameter nonlinear model on height measures of 13 boys and 14 girls from the French auxology survey and found evidence for midgrowth spurt in less than half of the children.

The extended splines fit growth curve before age 9 years by $\Phi_3(t;0,1,3)$ and velocity curve by $\Phi_2(t;-1,2)$, which allows the estimation of up to two inflection points for this period. Our individual velocity curves of height show small upward turn, not an obvious spurt, at about ages 5 to 10 years in 47 of 89 males and 13 of 67 girls by using eight-parameter extended splines for separate individuals. Our mean velocity curve of height does not show the upward turn in either gender.

It is not surprising that mean velocity curves do not exhibit the midgrowth spurt because of the variation of the timing and small intensity of these multiple minor fluctuations. The ages at take-off of the midgrowth spurt are spread over about four years with mean at about 6 years (Berkey, Reed and Valadian, 1983; Gasser, Müller, Köhler, Prader, Largo and Molinari, 1985). Gasser, Müller, Köhler, Molinari and Prader (1984) have noted that the mid-growth spurt is often small, almost drowned in random noise. Furthermore, some individuals can have more than one small spurt during childhood (Gasser, Kneip, Ziegler, Largo and Prader, 1990). Butler, McKie and Ratcliffe (1990) demonstrated the cyclical nature of prepubertal growth in the data of the Edinburgh Longitudinal Growth Study: a pre-school spurt at ages 4.8 and 4.6 years for boys and girls, midgrowth spurt at ages 7.0 and 6.7 years and a late-childhood spurt at 9.2 and 8.6 years.

A cyclical pattern may exist throughout growth and small spurts are more easily located at ages when growth is slower than when it is more rapid. These small spurts vary considerably both in size and timing and require an excessive increase in the number of parameters to describe. The extended splines with a few parameters cannot describe the pattern and the variation appropriately. Jolicoeur, Pontier and Abidi (1992) discuss this matter and point out that it is debatable whether the midgrowth spurt should necessarily be incorporated into a growth model of stature or whether it should be treated as an ignorable fluctuation among others.

5.7 Asymptotic Growth and Prediction

The extended splines are flexible but are not a satisfactory model for data with asymptotic adult values. A fractional polynomial with $p \le 1$ may have a less marked end effect than quadratic and cubic polynomials. As the height data in this study include adult measurements a fractional polynomial with $p \le 1$ is used in the last segment of the extended spline. The fractional polynomial of $\Phi_2(t; -1, 1)$ or $\Phi_2(t; -2, 1)$ was considered and the reciprocal terms were not significant and finally the $\Phi_2(t;0, 1)$ is chosen, which gives a slight overestimate of velocity at the end.

The presence of edge effects is common in curve fitting and smoothing techniques (Cole and Green, 1992). This problem may be partly due to the variation of height after the puberty spurt though height measurements are expected to asymptote to adult values at about 17 years for girls and 18 for boys. The experience of using the kernel estimation for individual height curves reveals that it is common to find a small bump after the pubertal spurt at ages 16 to 19 years (Gasser, Müller, Köhler, Molinari and Prader, 1984; Gasser, Kneip, Ziegler, Largo and Prader, 1990). In addition, the relative small number of measurements during this age range may affect the results. We would suggest including measurements up to 20 years, if possible, to obtain better results.

In general, extended splines are satisfactory for modelling growth. They are useful for purposes of description or comparison. However, similar to polynomials they can not safely be used to predict growth outside the delimited region as nonlinear asymptotic models do

(Bock 1989). But Goldstein (1989) describes an alternative approach to prediction using a multivariate multilevel model which can be used with our models. It can be expressed by including adult height measurements to the extended model (3.5):

$$y_{ij} = \beta_{0j} + \beta_{1j}t_{ij} + \beta_{2j}\ln(t_{ij}) + \beta_{3j}(9 - t_{ij})_+^3 + \beta_{4j}(11 - t_{ij})_+^3 + \beta_{5j}(13 - t_{ij})_+^3 + \beta_{6j}(15 - t_{ij})_+^3 + \beta_{7j}(17 - t_{ij})_+^3 + e_{ij}$$

$$y_j = \phi_0 + \alpha_j$$

where the first line of the model is for repeated measures of height and the second line for a single measure of adult height. All the parameters can be estimated by a 3-level model, where level 1 is the variable, repeated measures of height or adult height, level 2 is the occasion and level 3 is the individual. Other variables such as parent height or bone age can also be included into the model as illustrated by Goldstein (1989).

5.8 Further Work

Analysing data when a transformation required

In this study the work is focused on human growth in length not only because of its interest but also of its growth regularity. We may assume that most measurements in length have an approximately Normal distribution (Healy, Rasbash and Yang, 1988). However, to analyse human growth in weight it is necessary to transform data to approximate Normality. Logarithmic or Box-Cox transformations (Cole and Green, 1992) may be useful for this purpose. A transformation function on age may be required and the analysis may become more complicated when the data cover a wide age range. It would be interesting to study such transformations.

Another kind of transformation may be required in longitudinal studies on cognitive growth and development, where different scales are used over age. Plewis (1994) gives a review and discussion on statistical methods for understanding cognitive growth. He points out that different conclusions can be reached depending upon the linear transformations chosen.

Nonlinear Model

It is a linear modelling problem when knots of splines are given. It will become a non-linear problem when the knots are treated as continuously varying in a given range, that is, splines with variable knots. At present it is not clear whether it is worth trying to incorporate splines with variable knots into multilevel models.

It may be worth incorporating other nonlinear growth models into multilevel models. The triple logistic model of Bock and Thissen (1976) and the JPA2 model of Jolicoeur, Pontier and Abidi (1992) are often used for data on height from early childhood to adulthood. The nonlinear model of Roche, Mukherjee and Guo (1986) was used for head circumference. Goldstein (1995) describes a procedure for nonlinear models and gives examples with a single linear component, namely the Jenss model. In principle, the procedure can be applied for a model with multiple components, for example the triple logistic model with three components. In practice it is quite difficult for the procedure to reach convergence in fitting a model with multiple components.

Time series

We feel that it is reasonable to assume uncorrelated residuals in our model. Most of the data used in this study were collected at about 6 month intervals. A runs test on residuals, using SPSS/PC+ (version 3.0), for each individual was significant in 6 of 89 boys and 5 of 67 girls in height; in head circumference it was significant in 10 of 83 boys and 7 of 60 girls.

To make use of standard procedures for estimation of autocorrelation (Fuller, 1976, p236; Bock and Thissen, 1980, p274), it is necessary that the observations be equally spaced. The estimates are biased when the mean of the calculated residuals is used. Berkey (1982b) and Berkey, Reed and Valadian (1983) used the standard estimation of autocorrelation for the subset of the residuals which fall at approximately six-month intervals.

Autocorrelations (see Bock and Thissen, 1980, p274) calculated from the subset of the OLS residuals which occur at approximately 6 month intervals are given in Table 5.8.1. The results should be viewed with caution as there is considerable variability in the observed ages about target ages and a considerable number of missing measurements. The residual autocorrelations are relatively small compared with those found by Bock and Thissen (1980) in the triple logistic model. These autocorrelations and the runs test suggest that for these time lags, residuals can be assumed to be approximately independent. Many authors (El Lozy 1978; Preece and Baines, 1978; Berkey, 1982a, 1982b) make the same assumption.

Table 5.8.1 OLS residual autocorrelations at time lag (years)

Measures	Gender	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
HT	M	0.24	-0.07	-0.41	-0.38	-0.27	-0.02	-0.07	0.16
HT	F	0.15	-0.07	-0.33	-0.34	-0.26	-0.08	-0.00	0.10
HC	M	0.21	0.01	-0.11	-0.12	-0.17	-0.14	-0.16	-0.12
HC	F	0.24	0.04	-0.14	-0.15	-0.11	-0.17	-0.16	-0.07

Autocorrelations have also been calculated for the level 1 residuals of the random coefficient models in section 4. Table 5.8.2 shows that the autocorrelations estimated from the subset of residuals occur at approximately six month intervals. The residual autocorrelations are quite close to those of OLS in head circumference. The autocorrelations of lag 0.5, 1.0 and 1.5 of height measurements seem larger than those of found in OLS residuals, however they are close to those found by Bock and Thissen (1980) in triple logistic model.

Bock (1992) used triple logistic model on the Fels data gathered at 6-monthly intervals. The estimates of the means and covariances of his random effect model, when uncorrelated residuals are assumed, are essentially the same when autocorrelation are considered. Bock (1992) points out that if the data points are equally, and not too closely, spaced over the age range of interest, there is little harm in ignoring the autocorrelation. This is confirmed by other authors, for example, Goldstein, Healy and Rasbash (1994) who point out that the 6-monthly and yearly measurements can be regarded as independently varying about the growth curve with a constant variance that can be estimated from a suitable sample. We should keep in mind that large autocorrelations could also arise from misfit of the model. For example with a small sample, only a limited number of variables can be random at level 2 as we have discussed in the section 5.5 and section 4.2.

Table 5.8.2 Level 1 residual autocorrelations at time lag (years)

Measures	Gender	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
HT	M	0.60	0.28	-0.14	-0.34	-0.42	-0.38	-0.29	-0.19
HT	F	0.64	0.38	0.11	-0.10	-0.25	-0.31	-0.35	-0.31
HC	M	0.20	0.01	-0.10	-0.11	-0.15	-0.12	-0.14	-0.10
HC	F	0.28	0.07	-0.08	-0.11	-0.09	-0.15	-0.14	-0.06

Multilevel models allow us to analyse correlated residuals. Goldstein, Healy and Rasbash (1994) point out that if measurements are taken close enough together then their deviations from the fitted smooth curve are bound to be correlated. For height measurements on children prior to adolescence the point at which this 'autocorrelation' becomes apparent is for measurements made about 3 months apart. Goldstein, Healy and Rasbash (1994) proposed a time series model which deals with an autocorrelation for level 1 residuals for longitudinal data. First-and second-order autoregressive models are used to model level 1 residuals with a seasonal component in fixed part and also a model in continuous time is illustrated. Their time series model can be applied to our extended splines if autocorrelations are considered to be modelled. This is an interesting research topic for the future.

Finally and very importantly, the method of this thesis need to be extended to other measurements, both singly and jointly.

References

Berkey, C. S., 1982a, Bayesian approach for a nonlinear growth model. Biometrics, 38, 953-961.

Berkey, C. S., 1982b, Comparison of two longitudinal growth models for preschool children. Biometrics, 38, 221-234.

Berkey, C. and Laird, N. M., 1986, Nonlinear growth curve analysis: estimating the population parameters. Annals of Human Biology, 13, 111-128.

Berkey, C. S., Reed, R. B. and Valadian, I., 1983, Midgrowth spurt in height of Boston children. Annals of Human Biology, 10, 25-30.

Berkey, C. S. and Reed, R. B., 1987, A model for describing normal and abnormal growth in early childhood. Human Biology, 59, 973-987.

Berkey, C. S., Laird, N. M., Valadian, I. and Gardner, J., 1989, The analysis of longitudinal growth data with covariates. In Auxology 88: Perspectives in the science of growth and development, edited by J. M. Tanner (London: Smith-Gordon), pp.31-39.

Bock, R. D., 1989, Measurement of human variation: A two-stage model. In Multilevel analysis of education data, edited by D. Bock (New York: Academic Press), pp.319-342.

Bock. R. D., 1992, Structural and nonstructural analysis of multiphasic growth (personal communication).

Bock, R. D. and Thissen, D., 1976, Fitting multicomponent models for growth in stature. Proceedings of the 9th International Biometrics Conference (Boston) 1, 431-442.

Bock, R. D. and Thissen, D., 1980, Statistical problems of fitting individual growth curves. In Human physical growth and maturation, edited by F. E. Johnston, A. F. Roche and C. Susanne (New York: Plenum).

Bock, R. D., Toit, S. H. C. and Thissen, D., 1994, AUXAL: Auxological analysis of longitudinal measurements of human stature (Chicage: Scientific Software International).

Bock, R. D., Wainer, H., Peterson, A., Thissen, D., Murray, J., and Roche, A., 1973, A parameterization for individual human growth curves. Human Biology, 45, 63-80.

Bryk, A. S. and Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. Psychological Bulletin, 101, 147-158.

Bryk, A. S., Raudenbush, S. W. and Congdon, R. T., 1993, HLM2 and HLM3 Computer programs and users' guide, version 3.0 (Chicage: University of Chicago).

Buse, A. and Lim, L., 1977, Cubic splines as a special case of restricted least squares. Journal of the American Statistical Association, 72, 64-68.

Butler, G. E., McKie, M. and Ratcliffe, S. G., 1990, The cyclical nature of prepubertal growth. Annals of Human Biology, 17, 177-198.

Cole, T. J. and Green, P. J., 1992, Smoothing reference centile curves: the LMS method and penalized likelihood. Statistics in Medicine, 11, 1305-1319.

Count, E., 1943, Growth patterns of the human physique. Human Biology, 15, 1-32.

Cox, M. G., 1971, Curve fitting with piecewise polynomials. Journal of the Institute of Maths. and its Applications, 8, 36-52.

Deming, J., 1957, Application of the Gompertz curve to the observed pattern of growth in length of 48 individual boys and girls during the adolescent cycle of growth. Human Biology, 29, 83-122.

Dempster, A. P., Laird, N. M., and Rubin, D. B., 1977, Maximum likelihood with incomplete data via the EM algorithm (with discussion). Journal of the Royal Statistical Society, B, 39, 1-38.

Ertel, J. E., and Fowlkes, E. B., 1976, Some algorithms for linear spline and piecewise multiple linear regression. Journal of the American Statistical Association, 71, 640-648.

Et Lozy, M., 1978, A critical analysis of the double and triple logistic growth curves. Annals of Human Biology 5, 389-394.

Elston, R. C. and Grizzle, J. E., 1962, Estimation of time-response curves and their confidence bands. Biometrics, 148-159.

Eubank, R. L., 1984., Approximate regression models and splines. Communications in Statistics. Theory Methods A, 13, 433-484.

Fearn, T., 1975, A Bayesian approach to growth curves. Biometrika, 62, 89-100.

Fuller, W. A., 1976, Introduction to statistical time series. (New York: Wiley)

Gallant, A. R. and Fuller, W. A., 1973, Fitting segmented polynomial regression model whose join points have to be estimated. Journal of the American Statistical Association, 68, 144-147.

Gasser, T., Kneip, A., Binding, A., Prader, A. and Monlinari, L., 1991, The dynamics of linear growth in distance, velocity and acceleration. Annals of Human Biology, 18, 187-205.

Gasser, T., Kneip, A., Ziegler, P., Largo, R. and Prader, A., 1990, A method for determining the dynamics and intensity of average growth. Annals of Human Biology, 17, 459-474.

Gasser, T., Köhler, W., Müller, H-G., Kneip, A., Largo, R., Molinari, L. and Prader, A., 1984, Velocity and acceleration of height growth using kernel estimation. Annals of Human Biology, 11, 397-411.

Gasser, T., Müller, H-G., Köhler, W., Molinari, L. and Prader, A., 1984, Nonparametric regression analysis of growth curves. The Annals of Statistics, 12, 210-229.

Gasser, T., Müller, H-G., Köhler, W., Prader, A., Largo, R. and Molinari, L., 1985, An analysis of the mid-growth and adolescent spurts of height based on acceleration. Annals of Human Biology, 12, 129-148.

Gasser, T., Müller, H-G. and Mammitzsch, V., 1985, Kernels for nonparametric curve estimation. Journal of the Royal Statistical Society, B, 47, 238-252.

Gilks, W. R., Clayton, D. G., Spiegelhalter, D. J., Best, N. G., McNeil, A. J., Sharples, L. D. and Kirby, A. J., 1993, Modelling complexity: applications of Gibbs sampling in medicine. Journal of the Royal Statistical Society, B, 55, 39-52.

Goldstein, H., 1979, The design and analysis of longitudinal studies (London: Academic Press).

Goldstein, H., 1986a, Efficient statistical modelling of longitudinal data. Annals of Human Biology, 13, 129-141.

Goldstein, H., 1986b, Multilevel mixed linear model analysis using iterative generalized least squares. Biometrika, 73, 43-56.

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

Goldstein, H., 1989, Models for multilevel response variables with an application to growth curves. In Multilevel analysis of educational data, edited by D. Bock (New York, Academic Press), pp.108-125.

Goldstein, H., 1992, Efficient computational procedures for the estimation of parameters in multilevel models based on iterative generalised least squares. Computational Statistics & Data Analysis, 13, 63-71.

Goldstein, H., 1995, The multilevel analysis of growth data (submitted for publication).

Goldstein, H., Healy, M. J. R. and Rasbash, J., 1994, Multilevel time series models with applications to repeated measures data. Statistics in Medicine, 13, 1643-1655.

Guo, S. Siervogel, R. M., Roche, A. F. and Chumlea, W. M. C., 1992, Mathematical Modelling of Human Growth: A comparative study. American Journal of Human Biology, 4, 93-104.

Guo, S., Roche, A. F. and Moore, W. M., 1988, Reference data for head circumference and 1-month increments from 1 to 12 months of age. The Journal of Pediatrics, 113, 490-494.

Hauspie, R. C., Lindgren, G. W., Tanner, J. M. and Spruch, H. C., 1991, Modelling individual and average human growth data from childhood to adulthood. In Stability and Change: Models and Methods for Treatment of Data, edited by D. Magnusson, L. R. Bergman, G. Rudinger and B. Törestad (Cambridge: Cambridge University Press), pp.29-46.

Healy, M. J. R., 1989, Growth curves and growth standards-the state of the art. In Auxology 88. Perspectives in the science of growth and development, edited by J. M. Tanner (London: Smith-Gordon), pp.13-21.

Healy, M. J. R., Rasbash, J. and Yang, M., 1988, Distribution-free estimation of age-related centiles. Annals of Human Biology, 15, 17-22.

Jaffe, M., Tal, Y., Tirosh, E., Hadad, B. and Tamir, A., 1992, Variability in head circumference growth rate during the first 2 years of life. Pediatrics, 90, 190-192.

Jenss, R. M. and Bayley, N., 1937, A mathematical method for studing the growth of a child. Human Biology, 9, 556-563.

Jolicoeur, P., Pontier, J., Pernin, M.-O. and Sempe, M., 1988, A lifetime asymptotic growth curve for human height. Biometrics, 44, 995-1003.

Jolicoeur, P., Pontier, J. and Abidi, H., 1992, Asympototic Models for the Longitudinal Growth of Human Stature. American Journal of Human Biology, 4, 461-468.

Karlberg, J., 1989, A biologically-oriented mathematical model (ICP) for human growth. Acta Paediatr Suppl 350, 70-94.

Kreft, I. G. G., de Leeuw, J. and Kim, K-S., 1990, Comparing four different statistical packages for hierarchical linear regression: GENMOD, HLM, ML2 and VARCL. UCLA Statistics Series #50, 1-107.

Laird, N. M. and Ware, J. H., 1982, Random effects models for longitudinal data. Biometrics, 38, 963-974.

Largo, R. H., Gasser, T., Parader, A., Stutzle, W. and Huber, P. J., 1978, Analysis of the adolescent growth spurt using smoothing spline functions. Annals of Human Biology, 5, 421-34.

Leech, F. B. and Healy, M. J. R., 1959, The analysis of experiments on growth rate. Biometrics, 15, 98-106.

Lerman, P. M., 1980, Fitting segmented regression models by grid research. Applied Statistics., 29, 77-84.

Lindley, D. V. and Smith, A. F. M., 1972, Bayes estimates for the linear model (with discussion). Journal of the Royal Statistical Society, B, 34, 1-41.

Longford, N. T., 1987, A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects. Biometrika, 74, 817-827.

Marubini, G., Resele, L. F., Tanner, J. M. and Whitehouse, R. H, 1972, The fit of Gompertz and logistic curves to longitudinal data during adolescence on height, sitting height and biacromial diameter in boys and girls of the Harpenden growth study. Human Biology, 1972, 44, 511-524.

Meredith, H. V., 1981, An addendum on presence and absence of a mid-childhood spurt in somatic dimensions. Annals of Human Biology, 8, 473-476.

Merrell, M., 1937, The Relationahip of individual growth to average growth. Human Biology, 9, 37-70.

Paterson, L. and Goldstein, H., 1991, New statistical methods for analysing social structures: an introduction to multilevel models. British Educational Research Journal, 17, 387-393.

Plewis, I., 1985, Analysing Change: Measurement and Explanation Using Longitudinal Data (Chichester: Wiley).

Plewis, I., 1994, Statistical methods for understanding cognitive growth: a review, a synthesis and an application (submitted for publication).

Pfeffermann, D. and LaVange, L., 1989, Regression models for stratified multi-stage cluster samples. In Analysis of Complex Surveys, edited by C. J. Skinner, C. J., D. Holt and T. M. F. Smith (Chichester: Wiley), pp.237-260.

Poirier, D. J., 1973, Piecewise regression using cubic spline. Journal of the Americian Statistical Association, 68, 515-524.

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.

Preece, M. A., and Baines, M. J., 1978, A new family of mathematical models describing the human growth curves. Annals of Human Biology, 5, 1-24.

Prosser. R., Rasbash, J., and Goldstein, H., 1991, ML3-Software for three-level analysis: Users' guide (London: Institute of Education).

Rao, C. R., 1959, Some problems involving linear hypotheses in multivariate analysis. Biometrika, 46, 49-58.

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.

Ratcliffe, S.G., Masera, N., Pan, H. and McKie, M., 1994, Head circumference and IQ of children with sex chromosome abnormalities. Developmental Medicine and Child Neurology, 36, 533-544.

Ratcliffe, S. G., Pan, H. and McKie, M., 1992, Growth during puberty in XYY Boy. Annals of Human Biology, 19, 579-587.

Ratcliffe, S. G., Pan, H. and McKie, M., 1994, The growth of XXX females: population-based studies. Annals of Human Biology, 21, 57-66.

Ratcliffe, S. G. and Paul, N., 1986, Prospective studies in children with sex chromosome aneuploidy (New York: Alan. R. Liss).

Reed, R. B. and Berkey, C. S., 1989, Linear Statistical Model for Growth in Stature from Birth to Maturity. Americian Journal of Human Biology, 1, 257-262.

Roche, A. F., Mukherjee, D. and Guo, S., 1986, Head Circumference Growth Patterns: Birth to 18 Years. Human Biology, 58, 893-906.

Roche, A. F., Mukherjee, D., Guo, S. and Moore, W. M., 1987, Head circumference reference data: Birth to 18 years. Pediatrics, 79, 706-711.

Royston, P. and Altman, D. G., 1994, Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling. Applied Statistics, 43, 429-467.

Seber, G. A. F. and Wild, C. J., 1989, Nonlinear Regression (New York: Wiley), pp.433-459.

Silverman, B. W., 1985, Some aspects of the spline smoothing approach to non-parametric regression curve fitting (with Discussion). Journal of the Royal Statistical Society, B, 47. 1-52.

Smith, P. L., 1979, Splines as a useful and convenient statistical tool. The American Statistician, 33, 57-62.

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

Stützle, W., Gasser, Th., Molinari, L., Largo, R. H., Prader, A., and Huber, P.J., 1980, Shape-invariant modelling of human growth. Annals of Human Biology, 7, 507-28.

Tanner, J. M. and Cameron, N., 1980, Investigation of the mid-growth spurt in height, weight and limb circumferences in single-year velocity data from the London 1966-67 growth survey. Annals of Human Biology, 7, 565-577.

Tanner, J. M. and Whitehouse, R. H., 1976, Clinical longitudinal standards for height, weight, height velocity, weight velocity, and the stages of puberty. Archives of Disease in Childhood, 51, 170-179.

Tanner, J. M., Goldstein, H. and Whitehouse, R. H., 1970, Standards for children's height at ages 2 to 9 years, allowing for height of parents. Archives of Disease in Childhood, 45, 755-762.

Tanner, J. M., Whitehouse, R. H. and Takaishi, M., 1966, Standard from birth to maturity for height, weight, height velocity, and weight velocity: British Children, 1965, Part I and II. Archives of Disease in Childhood, 41, 454-471, 613-635.

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

Wishart, J., 1938, Growth-rate determinations in nutrition studies with the bacon pig, and their analysis. Biometrika, 30, 16-28.

Wegman, E. J. and Wright, I. W., 1983, Splines in Statistics. Journal of American Statistical Association, 78, 351-365.

Wold, S., 1974, Spline functions in data analysis. Technometrics, 16, 1-11.

Appendices

Appendix A Non-linear Models for Individual Height Curve

Triple Logistic Model (Bock and Thissen 1976; Bock, Toit and Thissen, 1994)

$$y = \frac{a_1}{1+e^{-b_1 t}} + \frac{a_2}{1+e^{-(b_2 t + c_2)}} + \frac{a_3}{1+e^{-(b_3 t + c_3)}},$$

where y is height at time t.

BTT Model (Bock, Toit and Thissen 1994)

$$y = \frac{a_1}{[1+e^{-b_1 x}]^{d_1}} + \frac{a_2}{[1+e^{-(b_2 x + c_2)}]^{d_2}} + \frac{a_3}{[1+e^{-(b_3 x + c_3)}]^{d_3}}.$$

The values of d_1, d_2 and d_3 are given, which were suggested to be 0.75, 0.75 and 1.2 individually in the manual of AUXAL program. Triple logistic model can be considered as BTT model with d_1, d_2 and d_3 are equal to 1.

JPA2 Model (Jolicoeur, Pontier and Abidi, 1992)

$$y = a\left\{1 - \frac{1}{1+[b_1(t+e)]^{c_1} + [b_2(t+e)]^{c_2} + [b_3(t+e)]^{c_3}}\right\}.$$

Preece-Baines Model 3 (Preece and Baines, 1978)

$$y = h_1 - \frac{4(h_1 - h_\theta)}{[e^{p_0(t-\theta)} + e^{p_1(t-\theta)}][1+e^{q_1(t-\theta)}]},$$

where p_0, p_1 and q_1 are rate constants, θ is a time constant, h_θ is height at $t = \theta$ and y is height at time t.

Appendix B Kernels for Nonparametric Curve Estimation

See Gasser, Muller, Kohler, Molinari and Prader (1984) for details.

For measurements $f_j(t_i)$ of subject j at time t_i we assume:

$$f_j(t_i) = f_j^*(t_i) + \varepsilon_j(t_i),$$

where $f_j^*(t_i)$ denotes true measurements at age t_i. Random variations $\varepsilon(t)$ are due to measurement error, seasonal effect, environmental conditions, etc.

Denote $\hat{f}(t)_\nu$ for $d^\nu f^*(t)/dt^\nu$ ($\nu = 0, 1, 2$) and determine it as:

$$\hat{f}_\nu(t) = \sum_{i=1}^{N} f(t_i)/g_i(t),$$

where the weight $g_i(t) = \frac{1}{b^{\nu+1}} \int_{s_{i-1}}^{s_i} W_\nu\left(\frac{t-x}{b}\right)dx,$

N is the number of measurement for an individual, b is bandwidth (smoothing parameter), $s_i = (t_{i+1} + t_i)/2$ and W_ν is weighting function for ν-th derivative. The weight function W_ν for $\nu = 0, 1, 2$ is:

$$\begin{array}{lll} W_0(u) = & \frac{15}{32}(7u^4 - 10u^2 + 3) & |u| \leq 1 \\ & 0 & |u| > 1 \\ W_1(u) = & \frac{105}{32}(-9u^5 + 14u^3 - 5u) & |u| \leq 1 \\ & 0 & |u| > 1 \\ W_2(u) = & \frac{315}{64}(77u^6 - 135u^4 + 63u^2 - 5) & |u| \leq 1 \\ & 0 & |u| > 1 \end{array}$$

Appendix C Maximum Likelihood Estimation Via IGLS

The Iterative Generalised Least Squares (IGLS) algorithm has been proposed for estimating the Γ and Σ of equation (2.8) with a complex level 1 covariance structure by Goldstein (1986, 1992). The IGLS procedure produces maximum likelihood (ML) estimates when the random variables have a multivariate normal distribution by minimizing

$$l(\Gamma, \Sigma| Y) = \ln|\Sigma| + (Y - XZ\Gamma)^T\Sigma^{-1}(Y - XZ\Gamma),$$

which is the loglikelihood function for Γ and Σ given Y under the assumption of multivariate normality.

The IGLS algorithm applied has two parts cycle (suppose at step k):

(a) obtaining $\hat{\Gamma}^{(k)}$ using

$$\hat{\Gamma}^k = ((XZ)^T(\hat{\Sigma}^{(k-1)})^{-1}(XZ)^T)^{-1}(XZ)^T(\hat{\Sigma}^{(k-1)})^{-1}Y$$

(b) Estimating the random parameters of $\hat{\Omega}_{(2)}^{(k)}$ and $\hat{\Omega}_{(1)}^{(k)}$ using $(Y - (XZ)\hat{\Gamma}^{(k)})(Y - (XZ)\hat{\Gamma}^{(k)})^T$, $X_{(2)}$ and $X_{(1)}$.

Following additional notation is introduced to express the computing of part (b).

Let Y^{**} be $[Y_1^{**T}, \ldots, Y_j^{**T}]^T$ while Y_j^{**} be $\mathrm{vec}((Y_j - (XZ)_j\hat{\Gamma})(Y_j - (XZ)_j\hat{\Gamma})^T)$. Let Γ^{**} be $[[\mathrm{vec}\Omega_{(2)}]^T, [\mathrm{vec}\Omega_{(1)}]^T]^T$ and X^{**} be the design matrix $\left[X_1^{**T}, \ldots, X_J^{**T}\right]$ relating Y^{**} to the random parameters (see Browne, 1974). Note the vec operator stacks the columns of a matrix from left to right to form a vector. Estimates of random parameters in Γ^{**} can be obtained using GLS method to the following regression

$$Y^{**} = X^{**}\Gamma^{**} + \zeta$$

where ζ is a vector of residuals and its covariance matrix denoted by Σ^{**}. Elements of $\hat{\Sigma}^{(m-1)}$ are used to obtain $\hat{\Sigma}^{**(m)}$ which in turn is used to estimate $\Gamma^{**(m+1)}$.

The procedure of the IGLS cycle of (a) and (b) is: firstly, to obtain the starting estimates of $\Sigma^{(0)}$ using an ordinary least squares regression by setting $\Sigma^{(0)} = \sigma^2 I$; and then to update it successively until it satisfies a convergence criterion.

The final estimates of Σ and Σ^{**} from the above IGLS cycle are used to compute the variance and covariance of $\hat{\Gamma}$ and $\hat{\Gamma}^{**}$ as follows:

$$((XZ)^T\hat{\Sigma}^{-1}(XZ))^{-1},$$

$$2(X^{**T}\hat{\Sigma}^{**-1}X^{**})^{-1}$$

respectively. The estimation of IGLS at each iteration is consistent and the consistency does not depend on Y having a multivariate normal distribution (Goldstein, 1986b).