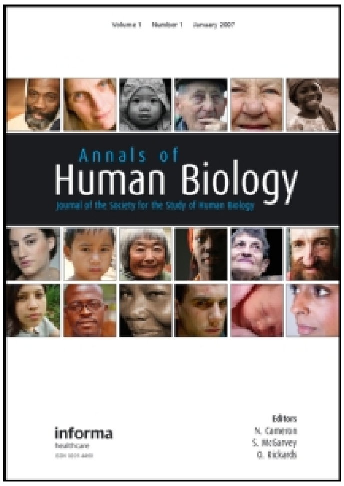


Annals of Human Biology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713723502

Shape-invariant modelling of human growth

W. Stützle [a]; Th. Gasser [b]; L. Molinari [c]; R. H. Largo [c]; A. Prader [c]; P. J. Huber [d]
[a] Department of Statistics, Stanford University, [b] Zentralinstitut für Seelische Gesundheit, Mannheim, FR Germany [c] Department of Paediatrics, University of Zurich, [d] Department of Statistics, Harvard University,

Online Publication Date: 01 December 1980

To cite this Article 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,7:6,507 — 528

To link to this Article: DOI: 10.1080/03014468000004641

URL: http://dx.doi.org/10.1080/03014468000004641

ANNALS OF HUMAN BIOLOGY, 1980, VOL. 7, NO. 6, 507–528

Shape-invariant modelling of human growth

W. STÜTZLE
Department of Statistics, Stanford University

TH. GASSER
Zentralinstitut für Seelische Gesundheit, Mannheim, FR Germany

L. MOLINARI, R. H. LARGO and A. PRADER
Department of Paediatrics, University of Zurich

and P. J. HUBER
Department of Statistics, Harvard University

Received 15 June 1978; revised 22 July 1980

Summary. A new approach to modelling human height growth is presented which is also suitable for other variables. In a mathematical algorithm, some guess about the functional form of this growth process is improved consistently using the data; the resulting shape-invariant model (SIM) allows an approximately bias-free fitting for longitudinal data from 1 to 20 years with six parameters assigned to each individual. The SIM approach and the use of velocities rather than distances proved to be suitable for biomathematical modelling in order to answer questions qualitative in nature. In comparison of an additive two-component SIM and one where the appearance of puberty inhibits further growth of the non-pubertal component ('switch-off model'), the latter proved to be superior in various aspects. The model structure turned out to be the same for boys and girls, both before and during puberty. Among the qualitative features found are notable: a pronounced midspurt, and a dip before the onset of puberty, as well as the asymmetry of the pubertal peak. A preliminary analysis of individual parameters confirmed some results found previously by other methods regarding sex differences and relations between parameters and with adult height.

1. Introduction

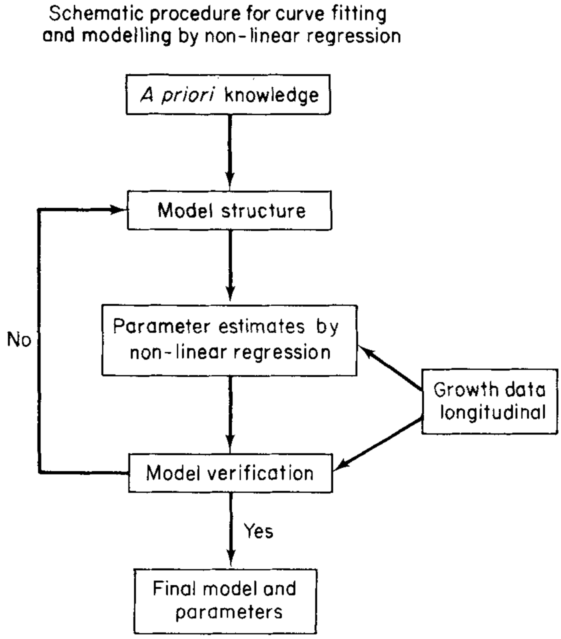
Cross-sectional data, or cohort studies, yield clinically valuable information for many problems of growth and development. Longitudinal growth data offer a unique chance for studying the dynamics of growth as a phenomenological correlate of physiological and in particular of endocrinological mechanisms. A detailed analysis of the variable height should lead to future benefits by laying a firm methodological basis for further variables and by providing parameters of height growth to be related to characteristics of pubertal and bone development. The importance of assigning a set of individual parameters to each growth curve has repeatedly been emphasized (Preece 1978, Largo *et al.* 1978). This goal can be achieved with reasonable accuracy by postulating some carefully selected model structure and performing non-linear regression ('fitting approach'; Marubini *et al.* 1971, Preece and Baines 1978), or else by extracting characteristic values from non-parametric regression curves ('smoothing approach'; Tanner *et al.* 1966, Largo *et al.* 1978).

To overcome the problem of prescribing somewhat arbitrarily a functional form for the growth process, we introduce the statistical concept of shape-invariant modelling (SIM); moreover, elements of biomathematical modelling are introduced by allowing both for the superposition and for the interaction of two components of growth

0301-4460/80/0706 0507 $02·00

velocity. We then obtain a parsimonious parametric fitting from 1 to 20 years of age without any appreciable bias.

For further discussion the following diagram may be useful:



For true biological modelling one usually needs input variables, e.g. hormone concentrations in our case. When only phenomenological quantities are available (height, weight etc.) one often has to fall back on descriptive models. The estimated residuals (= observed value − fitted value) provide measures of quality for competing descriptive models (model verification); this yields a basis for judging the validity of hypotheses about the dynamics of growth underlying these models. Consistency with biological knowledge, including pathological growth patterns, is of great importance. With the SIM approach, the iterative improvement of the model (see diagram) is performed within the algorithm: we take advantage of the availability of a rather large sample of individual curves, and estimate the model structure and the parameters for individuals iteratively from the data (with a fixed model, measurements of one curve only are used at a time). The estimated model reveals interesting features of the growth process: a midspurt at about 7 years, a dip in velocity before puberty and a marked asymetry of the pubertal peak.

It is not clear *a priori* whether one and the same model is appropriate for all individuals. Here we check whether the same model structure is adequate for boys and for girls.

Phenomenologically we are inclined to model the growth process from birth to adulthood with two components, one associated with pubertal and the other with non-pubertal growth. It is mathematically simplest to postulate additivity, but an interaction of the two components is perhaps biologically more realistic: the appearance of pubertal growth would smoothly terminate prepubertal growth ('switch-off model'). We present evidence that the switch-off model is more consistent with the observations than an additive model.

In this contribution we focus on qualitative aspects, associated with the model structure, whereas the statistical analysis of the individual parameters (quantitative aspects) is only a preliminary one.

It seems that the SIM method might easily generalize to other variables than height, while this is more questionable for fixed models so far discussed in the literature.

2. Subjects and methods

Subjects

In 1954 a prospective longitudinal study of growth and development of 413 healthy Swiss children was initiated at the Kinderspital Zürich, in coordination with four other longitudinal studies (Brussels, London, Paris, Stockholm), and the Centre International de l'Enfance in Paris (Falkner 1961).

In this report, children are excluded from the analysis if height measurements between birth and 3 years are incomplete, or if later more than three or two consecutive observations are missing. None of the children included was suffering from a disease hampering growth; six boys and two girls of very tall stature, treated with sex hormones, were excluded. From the remaining children random samples of 45 boys and 45 girls were selected for statistical analysis. For cross-validatory purposes these samples were subdivided at random into three samples of 15 boys and three samples of 15 girls; a recombination yielded three random samples each with $N = 30$ of mixed sex.

Measurements

Before the age of 9 years in girls, and 10 years in boys, the children were measured annually at birthday $\pm$ 14 days. Afterwards they were measured six-monthly within the same limits until the annual increment in height was less than 0·5 cm/year. Thereafter, the measurements were done again annually and were discontinued when the increment had become less than 0·5 cm in 2 years.

Editing

Gross errors had been corrected in a previous investigation, and missing observations were completed to simplify computations as described there (Largo *et al.* 1978).

Computation

The computations were performed on a PDP-10 computer. The interactive mode of operation and the availability of graphics facilities proved to be important (Tektronix 4014, Benson plotter). The program for shape-invariant modelling as well as the application graphics software is our own development (the non-linear least squares module was later exchanged for one written at MIT; Dennis *et al.* 1978). For one girl the algorithm yielded nonsense parameters for different initial values and she, therefore, was not included in the further analysis.

Mathematical model

Notation:

$h_i(t_j)$ = observed height of individual i ($i = 1, \ldots, N$) at age t_j ($j = 1, \ldots, n$)

$v_i[(t_{j+1} + t_j)/2] = [h_i(t_{j+1}) - h_i(t_j)]/(t_{j+1} - t_j)$ = 'observed' velocity at age $(t_{j+1} + t_j)/2$

$H_i(t)$ $[\hat{H}_i(t)]$ = true [estimated] height of individual i at age t

$V_i(t)$ $[\hat{V}_i(t)]$ = true [estimated] velocity of individual i at age t

$t_j^* = (t_{j+1} + t_j)/2$ = ages for velocity ($j = 1, \ldots, n-1$)

The true height $H_i(t)$ is defined by equation (1). Since the function H_i is unknown, the residuals ε_{ij} in reality contain a bias depending on the choice of H_i; for the ease of notation this dependence is suppressed.

In view of the difficulty in proposing an adequate model structure, a number of authors have restricted themselves to a limited age span, in particular the adolescent period (Deming 1957, Marubini *et al.* 1971, Tanner *et al.* 1976); most of the functional models suggested can then be summarized in the following general regression equation:

$$h_i(t_j) = H_i(t_j) + \varepsilon_{ij} \qquad (1)$$

$$= a_i S\left(\frac{t_j - b_i}{c_i}\right) + d_i + \varepsilon_{ij},$$

where S = shape function; examples are logistic function $S_L(x) = 1/[1 + \exp(-x)]$ and Gompertz function $S_G(x) = \exp[-\exp(-x)]$;

a_i, b_i, c_i = individual parameter for the intensity, time and duration of the pubertal spurt;

d_i = height at some age before puberty;

ε_{ij} = residual term, comprising the measurement error and individual random fluctuations of growth, and the bias, if the postulated model S is not correct ($E(\varepsilon_{ij}) \neq 0$); the ε_{ij} are assumed to be stochastically independent for all i and j, $\mathrm{Var}(\varepsilon_{ij}) = \sigma_{ij}{}^2$ ($= \sigma_i^2$ if homoscedastic).

In this paper, the term 'shape-invariant modelling' is used for estimating the best-fitting function S from the data—among all functions satisfying some smoothness requirements—instead of restricting us to one or a few *a priori* fixed functions. Moreover, we are looking for a model fitting well over the whole growth process, and to achieve this, we will consider the two-component models (4), (5), with six parameters each. Preece and Baines (1978) have suggested three new models, among which model 3 (also with six parameters) proved to be particularly successful in fitting height growth from 4 years onwards.

$$H_i(t) = H_{a,i} - \frac{4(H_{a,i} - H_{\theta,i})}{\exp p_{0,i}(t-\theta_i) + \exp p_{1,i}(t-\theta_i)[1 + \exp q_{1,i}(t-\theta_i)]}, \qquad (2)$$

where

$$H_a = \text{adult height}, \quad H_\theta = H(\theta)$$

An earlier attempt at a global fitting by Bock *et al.* (1973) (1–18 years) had led to high residual mean square errors, indicating severe lack of fit, and to other discrepancies; their model consisted in the addition of two logistic functions, a two-component additive shape-invariant model in our terminology:

$$H_i(t) = a_{1i} S_L\left(\frac{t - b_{1i}}{c_{1i}}\right) + (H_{a,i} - a_{1i}) S_L\left(\frac{t - b_{2i}}{c_{2i}}\right) \qquad (3)$$

Let us now delineate our approach (for more details see Stützle, (1977) and also Lawton *et al.* (1972), who first suggested such a procedure in a different context). As in Largo *et al.* (1978), we start from height velocities instead of the customarily used distance curves. El Lozy (1978) has convincingly illustrated how bad a biased estimate can look in terms of velocity. Even more important is that velocities are closer to the dynamics of growth and therefore better suited for biomathematical modelling (compare model (5) below). These advantages well outweigh the disadvantages, i.e. that velocities are not observed but merely approximated from distances, and that even if residuals of distances ε_{ij} were uncorrelated and homoscedastic across age, the residuals of velocities e_{ij} will become correlated and heteroscedastic within individuals. How these deviations from classical regression are treated is described in an appendix.

Two different two-component SIMs are compared, one additive:

$$v_i(t_j^*) = a_{1i} S_1\left(\frac{t_j^* - b_{1i}}{c_{1i}}\right) + a_{2i} S_2\left(\frac{t_j^* - b_{2i}}{c_{2i}}\right) + e_{ij}, \tag{4}$$

while in the other ('switch-off model'), further growth due to the first component is inhibited by the appearance of the second component (the same notation as in (4) is used to keep it simple):

$$v_i(t_j^*) = a_{1i} S_1\left(\frac{t_j^* - b_{1i}}{c_{1i}}\right)\varphi\left(\frac{t_j^* - b_{2i}}{c_{2i}}\right) + a_{2i} S_2\left(\frac{t_j^* - b_{2i}}{c_{2i}}\right) + e_{ij}. \tag{5}$$

It is tempting to associate the two components with pubertal (S_2) and non-pubertal growth (S_1 or $S_1.\varphi$); such an interpretation is meaningful for the switch-off model (5) (compare sections 3(a), (c)). The function φ—coupled in location and scale to the second component—has to decrease to zero, and it was chosen as a Gaussian ogive (compare the start function S_2^0 in (7):

$$\varphi(y) = 1 - \frac{\int_{-\infty}^{y} \exp(-t^2)\,dt}{\int_{-\infty}^{+\infty} \exp(-t^2)\,dt}. \tag{6}$$

Were S_1 and S_2 given, we would have one non-linear regression per individual, with six parameters ($a_{1i}, \ldots, c_{2i}$) to be estimated by minimizing the residual sum of squares (RSS). If, on the other hand, the individual parameters were known, we could estimate the shape functions from the aligned individual curves. Since neither is known, we have to determine the parameters and the shape-functions S_1, S_2 iteratively; start with some functions S_1^0, S_2^0 from previous experience (Largo *et al.* 1978):

$$\begin{aligned} S_1^0(x) &= 1 + \exp(-x), \\ S_2^0(y) &= \exp(-y^2). \end{aligned} \tag{7}$$

Following a suggestion of one referee of an earlier draft, we used alternatively the first derivative of the logistic function as start shape-functions (compare Bock *et al.* 1973, and equation (3)):

$$\begin{aligned} S_1^0(x) &= \frac{d}{dx} S_L(x) = \exp(-x)/[1 + \exp(-x)]^2, \\ S_2^0(y) &= \frac{d}{dy} S_L(y) = \exp(-y)/[1 + \exp(-y)]^2. \end{aligned} \tag{8}$$

To begin with, individual parameters are computed with respect to S_1^0, S_2^0 by non-linear least squares. Then the shape-functions S_1^0, S_2^0 are improved, yielding functions S_1^1, S_2^1; in this step it is important to choose the correction functions from a flexible class easy to handle computationally (linear least squares); our choice fell on B-splines, where the only arbitrariness lies in the number and position of knots. Now individual parameters are computed with respect to S_1^1, S_2^1, etc. This algorithm, yielding the optimal SIMs (4), (5) and the respective parameters, went back and forth from estimating individual parameters to improving at least five times before terminating according to some convergence criterion. For further details on models and methods, the reader is referred to the appendix. To search for developmental age scales—where children are at comparable developmental stages—is not new, and a partial answer is obtained by shifting the individual curves in such a way that the pubertal peaks coincide (Tanner *et al.* 1966; comparable to forming $(t-b_{2i})$ in formula (9)). Shape-invariant modelling allows a more formal and a more complete answer to this problem by introducing standardized age scales for child no. i at a chronological age t with respect both to the first $(x=x_i(t))$ and to the second $(=y_i(t))$ component:

$$x_i(t)=\frac{t-b_{1i}}{c_{1i}},\qquad y_i(t)=\frac{t-b_{2i}}{c_{2i}}. \tag{9}$$

The analysis of residuals is a powerful statistical tool for the detection of a model inadequacy. Since a model bias is related rather to some stage of development than to the chronological age, we introduce 'aligned residuals', i.e. residuals on standardized age scales, and this allows it to pool residuals across individuals (of both sexes even). Otherwise, bias phenomena would get blurred in the same way as the pubertal peak when averaging individual curves with respect to chronological age.

Our statistical analysis of height growth goes from 1 to 20 years (with $N=31$ height measurements). It is possible to include the first year, and this was done at first. Reasons for the exclusion later on were:

(*a*) The variability is rather high, due to both larger measurement errors and more influential environmental factors, in particular pregnancy (compare e.g. Smith *et al.* 1976).

(*b*) Our computer program is based on equidistant knots for the splines, and this may not be fully adequate for the rapid changes in the first year of life.

3. Results

(a) Switch-off and the additive model

In this section we check if the models defined by (4) and (5) provide a good fit to the data, and compare them; as a contrast, results for the differentiated double logistic function (8) (Bock *et al.* 1973) are also given. The measures of goodness-of-fit used were the following:

(i) Aligned residuals
—plotted against the standardized age scale;
—tested on expectation zero (=no bias) in segments of standardized age.
This was done using the whole sample, and not individually.

(ii) Residual mean squares (RMS)
RMS = sample variability + sample bias squared.

Regarding (ii), the estimated residuals of velocities e_{ij} are weighted in such a way that the variability (but not the bias) is comparable to that of distances (see appendix).

Table 1 summarizes the result of one-sample Wilcoxon tests applied to the estimated residuals of the switch-off model (5) aligned separately with respect to the first and to the second component. There are 20 age classes to check for an age-specific

Table 1. Goodness-of-fit for switch-off SIM-model ($N = 30$): Wilcoxon test on grouped aligned residuals.

	First component		Second component	
Age of average child (years)	N	Wilcoxon statistic standardized	N	Wilcoxon statistic standardized
0·0 – 0·4	2	0·00	10	−0·51
0·4 – 0·7	14	0·31	16	1·53
0·7 – 1·9	29	0·11	23	−0·68
1·9 – 3·0	32	−1·13	25	−0·97
3·0 – 4·2	37	−2·20*	30	−1·58
4·2 – 5·3	31	2·30*	35	−0·65
5·3 – 6·5	38	1·02	33	0·79
6·5 – 7·6	61	−0·46	36	0·51
7·6 – 8·8	37	0·95	34	1·71
8·8 – 9·9	30	1·34	35	−1·07
9·9 – 11·1	25	−2·39*	35	1·30
11·1 – 12·2	21	0·00	35	−0·37
12·2 – 13·4	23	−0·02	33	0·86
13·4 – 14·5	20	0·24	34	−1·13
14·5 – 15·7	20	−0·43	33	0·32
15·7 – 16·8	17	1·28	34	−0·60
16·8 – 18·0	17	−0·05	34	0·67
18·0 – 19·1	17	0·52	27	0·11
19·1 – 20·3	14	−0·75	11	0·40
20·3 – 21·4	16	1·22	6	0·00

Table 2. Goodness-of-fit for additive SIM-model ($N = 30$): Wilcoxon test on grouped aligned residuals.

	First component		Second component	
Age of average child (years)	N	Wilcoxon statistic standardized	N	Wilcoxon statistic standardized
0·0 – 0·4	15	−1·33	16	−2·35*
0·4 – 0·7	15	−0·43	16	1·11
0·7 – 1·9	12	1·14	21	1·46
1·9 – 3·0	20	0·95	26	0·18
3·0 – 4·2	28	1·13	27	−1·67
4·2 – 5·3	34	−0·51	30	0·60
5·3 – 6·5	33	0·00	32	−1·19
6·5 – 7·6	38	1·46	28	2·74**
7·6 – 8·8	31	−1·91	30	1·01
8·8 – 9·9	37	1·34	34	0·21
9·9 – 11·1	34	−0·94	33	0·77
11·1 – 12·2	34	1·49	36	−1·42
12·2 – 13·4	33	−0·91	34	0·89
13·4 – 14·5	32	−1·43	32	−1·60
14·5 – 15·7	29	−0·39	35	−0·35
15·7 – 16·8	29	0·22	31	0·24
16·8 – 18·0	22	−2·27*	28	−2·11*
18·0 – 19·1	22	−2·24*	22	−1·46
19·1 – 20·3	16	−0·13	17	−1·70
20·3 – 21·4	10	0·92	8	−1·89

Table 3. Goodness-of-fit statistic for double logistic model, velocity ($N = 30$): Wilcoxon test on grouped aligned residual.

	First component		Second component	
Age of average child (years)	N	Wilcoxon statistic standardized	N	Wilcoxon statistic standardized
0·0– 0·4	11	1·56	16	−0·34
0·4– 0·7	17	0·62	18	−1·61
0·7– 1·9	23	−0·20	22	−1·40
1·9– 3·0	24	−1·53	21	−0·10
3·0– 4·2	30	−1·34	25	−1·18
4·2– 5·3	30	0·49	24	1·50
5·3– 6·5	35	1·56	30	2·96**
6·5– 7·6	39	0·54	27	3·09**
7·6– 8·8	31	3·36***	30	3·70***
8·8– 9·9	33	2·51*	34	3·74***
9·9–11·1	33	2·18*	37	1·67
11·1–12·2	38	2·22*	32	−2·35*
12·2–13·4	31	−0·77	35	1·79
13·4–14·5	35	−1·40	34	−1·05
14·5–15·7	35	−3·14**	35	−4·94***
15·7–16·8	27	−3·12***	29	−3·27**
16·8–18·0	24	−4·16***	30	−4·73***
18·0–19·1	19	−3·80***	22	−3·96***
19·1–20·3	14	−3·26**	13	−3·07**
20·3–21·4	9	−2·25*	11	−2·89**

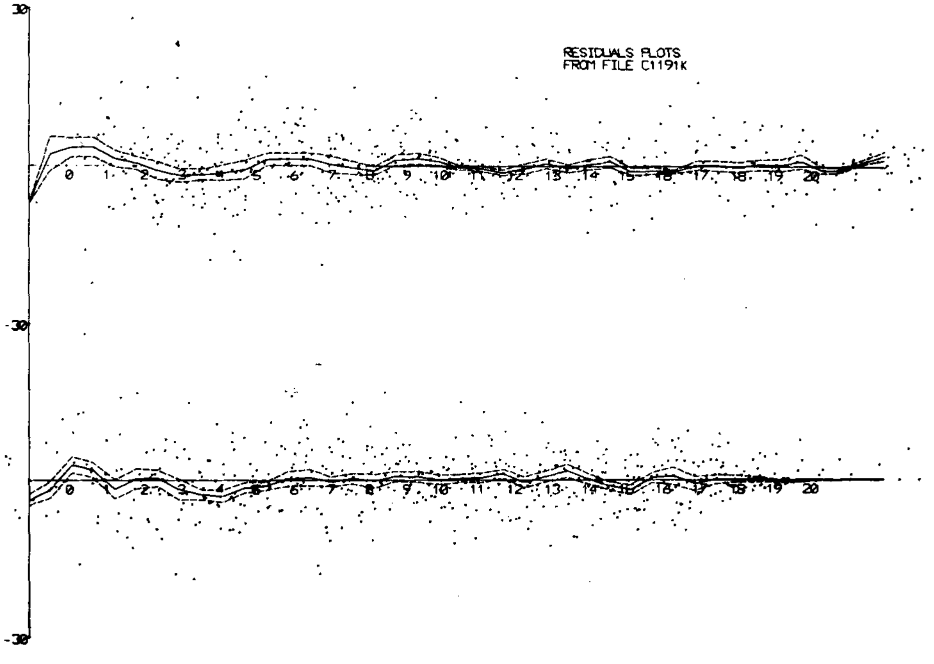


Figure 1. Residuals for switch-off SIM ($N = 30$, first random sample); above (below) aligned with respect to first (second) component. Heavy and dotted lines = running median ± standard error of the median estimated by the median deviation, standardized for the normal distribution.

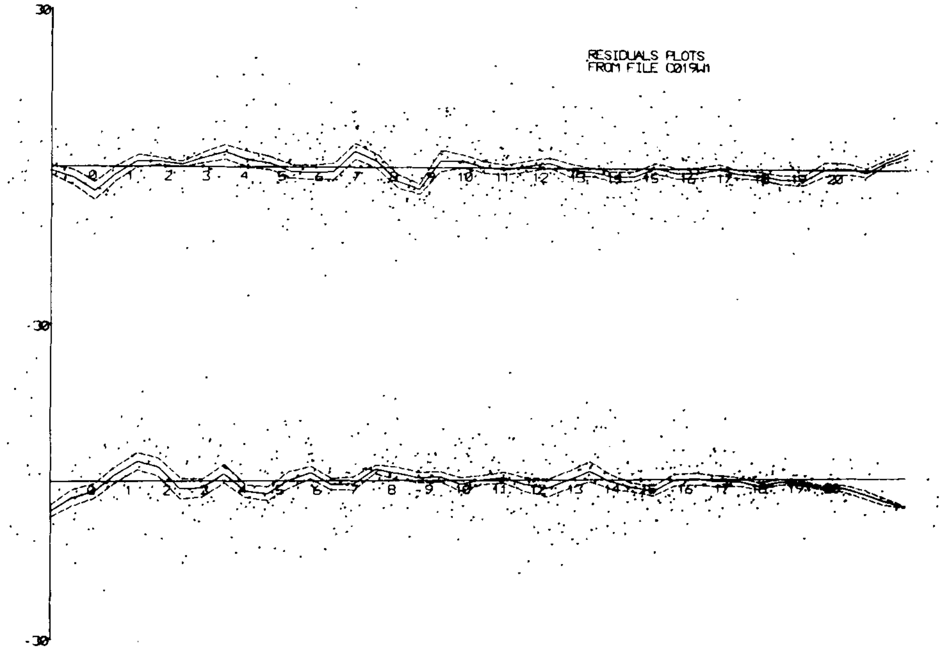


Figure 2. Residuals for additive SIM ($N=30$, first random sample); conventions as for figure 1.

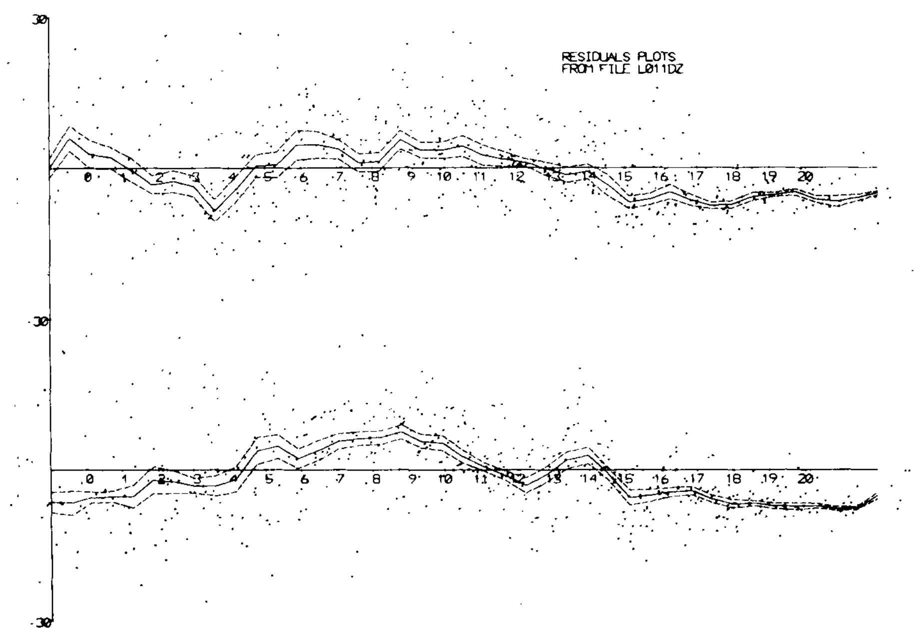


Figure 3. Residuals for derived double logistic model ($N=30$, first random sample); conventions as for figure 1.

bias) and *, **, *** indicate error probabilities $\leqslant 5\%$, $\leqslant 1\%$, $\leqslant 0{\cdot}1\%$ (with respect to the asymptotically valid normal distribution).

Table 2 contains the same information for the additive model (4). For both models the bias is negligible in magnitude, and there is no time trend for the switch-off model and a modest negative bias in the adolescent period for the additive model with respect to both components. When ranking all the values of tables 1 and 2 according to their absolute magnitude, 18 out of the 25 most deviant values have to be attributed to the additive model (4), which leads to some preference for the switch-off model. Table 3 demonstrates how bad the double logistic model is with respect to bias. The plots of the aligned residuals in figures 1–3 confirm these conclusions.

Table 4 allows a neat comparison (in terms of residual mean squares) of the properties of a fixed model versus an improved model and of an additive versus a switch-off model; it is notable that we win almost as much when introducing the

Table 4. Effects of including switch-off and of iteratively improving the ($N=30$) model on the RMS (in mm^2).

	Boys ($N=15$)			Girls ($N=15$)		
	Additive		Switch-off	Additive		Switch-off
Fixed model	38 DL	42 EXP	26 EXP	37 DL	37 EXP	24 EXP
Improved model	24 DL		21 EXP	22 DL		20 EXP

DL start model: double logistic differentiated.
EXP start model: $S_1(t)=1+\exp(-t)$, $S_2=\exp(-t^2)$.

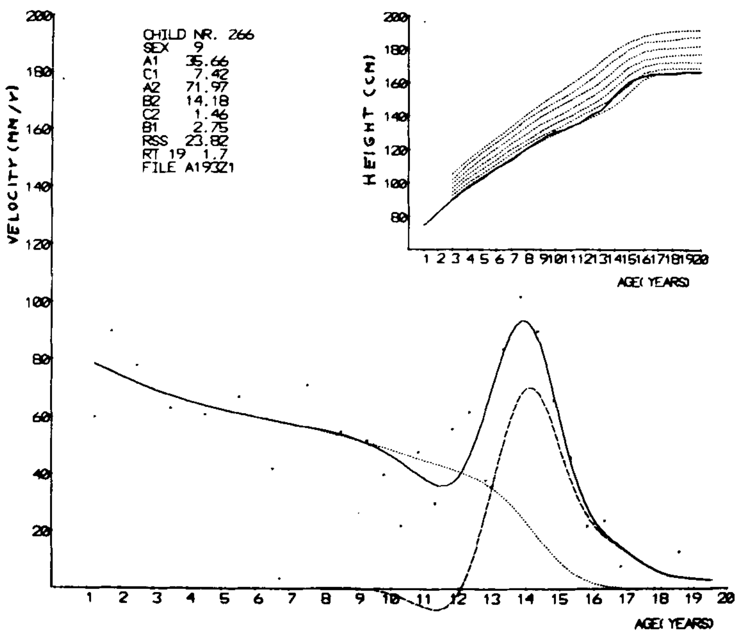


Figure 4. Small boy at the third centile of height distance (above); switch-off SIM fit, RMS $=23{\cdot}82$ mm^2.

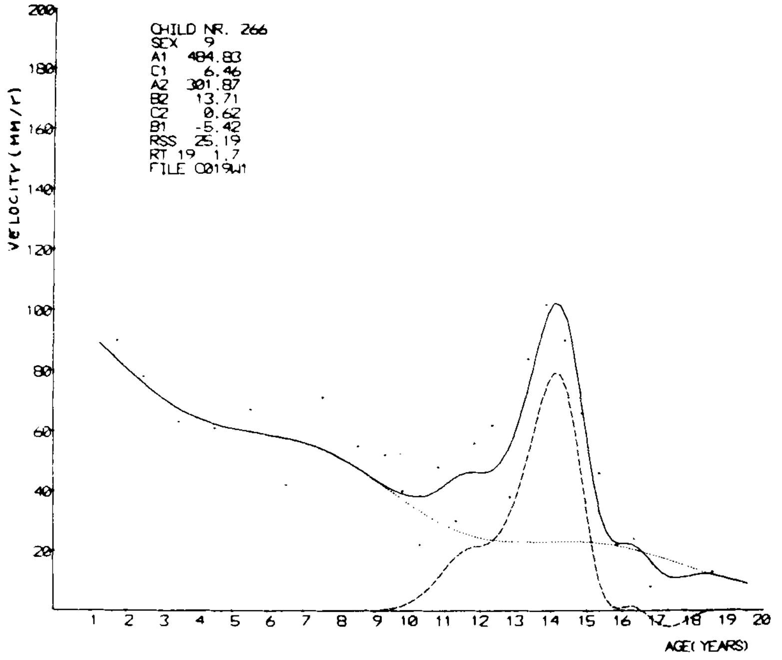


Figure 5. As in figure 4, but additive model, RMS$=25\cdot19\,\mathrm{mm}^2$.

biomathematical concept of a fixed two-component additive SIM. The combination of the two new concepts brings the best results.

The overall conclusions are the following: the data are in better consistence with a switch-off model than with an additive model. Whenever biomathematical concepts—such as the switch-off—are lacking, optimizing the model structure across the sample of curves may improve the fit appreciably compared to some *a priori* guess.

Let us illustrate the two models by two examples: the first is a small boy, who moves along the third centile of height distance (figures 4 and 5). The fit is satisfactory in both models as judged visually and by RMS. In the additive model the pubertal peak is superposed on the first component going on beyond puberty.

The second example is a late-maturing girl of average height at almost all ages (figures 6 and 7). From residual mean squares (RMS), one would conclude that the fit for the additive model ($17\cdot25\,\mathrm{mm}^2$) is at least as good as for the switch-off model ($20\cdot69\,\mathrm{mm}^2$). The composition of the two components is qualitatively the same as in the previous example for the switch-off model but not for the additive model: here the second component stretches well into the first component and vice versa. This overlapping of the two additive components can be seen in many children and also in the respective shape functions (compare, however, figures 9 and 10).

(*b*) *Shape functions for boys and girls*

In this and the next section we restrict ourselves to the switch-off model. It is well known that the quantitative features of height growth—the individual parameters—are different for boys and girls (Largo *et al.* 1978, Preece and Baines 1978). It is not clear whether this also applies to the model. The fact that the residual mean square errors reported in the literature were consistently larger for boys, and the endocrinological

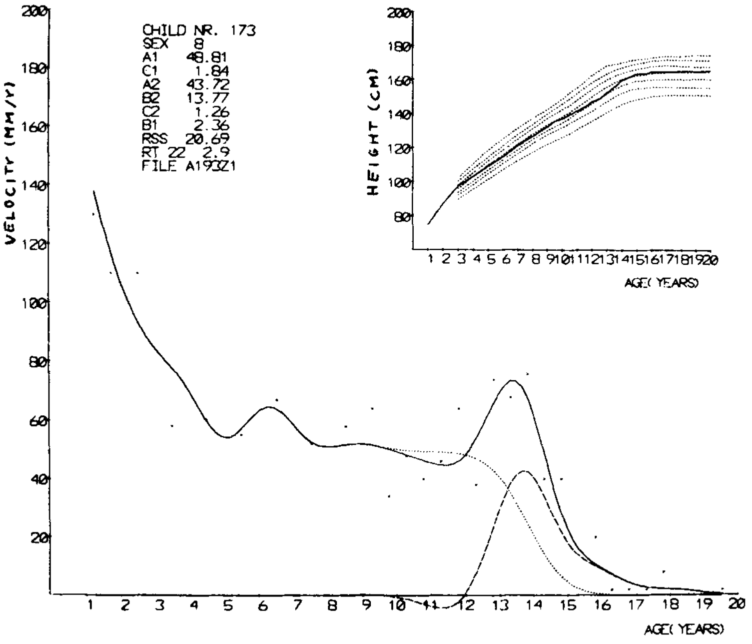


Figure 6. Tall girl at the 97 centile of height distance with later puberty (above), switch-off SIM f RMS = 20·69 mm^2.

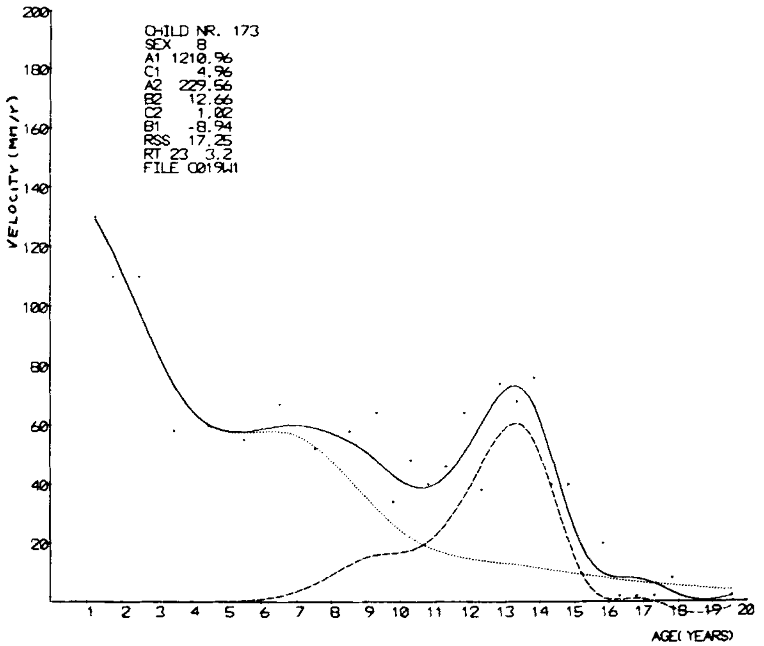


Figure 7. As in figure 6, but additive model, TMS = 17·25 mm^2.

difference seen in puberty underline the importance of this question. Our analysis, however, showed that the two shape functions are identical (up to random variations) for both sexes.

In table 5 functions estimated for 15 girls and for 15 boys separately, and for 30 children combined, are compared for the three random samples in terms of residual sum of squares (extra sum-of-squares principle). There is no tendency in the pseudo F-values for rejecting common shape functions for both sexes.

Table 5. One or two models for boys and girls: comparison of RSS for SIM with switch-off.

Sample	RSS ♂, 15	RSS ♀, 15	df	RSS_{30} (pooled)	df	$F_{14,662}$†
1	6907·8	6677·3	331	14094	676	1·77*
2	8247·6	4628·4	331	12919	676	0·16
3	8151·9	7682·1	331	15834	676	0·00

† Pseudo F-values.

(*c*) *Qualitative results*

Before discussing features contained in the shape functions, we have to check how well shape functions are determined from a sample of N curves in terms of variability and bias. It is to be expected that the variability depends primarily on the sample size, whereas the bias depends on the choice of classes for improving a start model, the number and position of knots for the splines in our case. As can be inferred from figure 8, the pattern of the shape functions is consistent for the three random samples of $N = 30$ children (figure 8, together with SIM correction functions); this is also true for

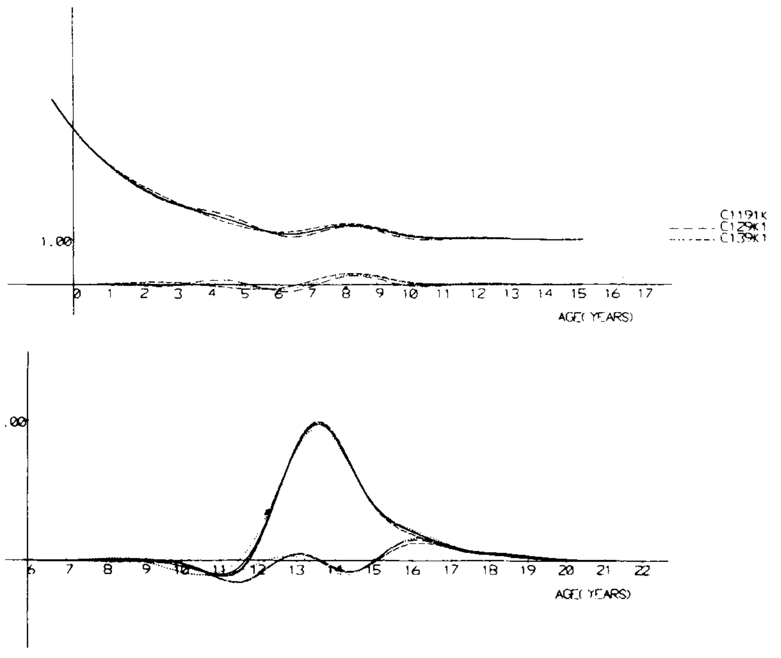


Figure 8. Shape functions for switch-off SIM first component above, second component below (with corrections). For three random samples with $N = 30$. Corrections for the start model (7) are given in the same graph.

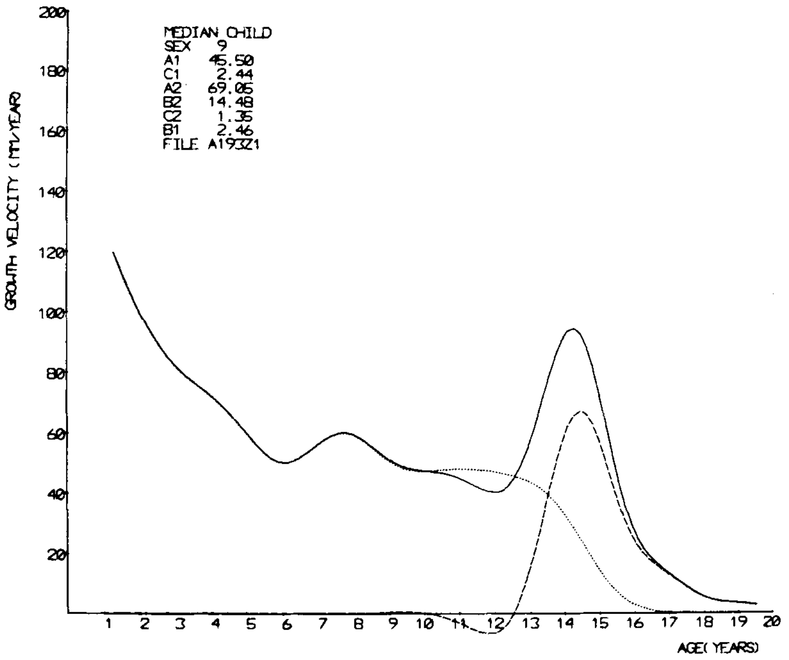


Figure 9. Growth velocity curve for median boy, i.e. with the median of the parameters inserted into the switch-off SIM.

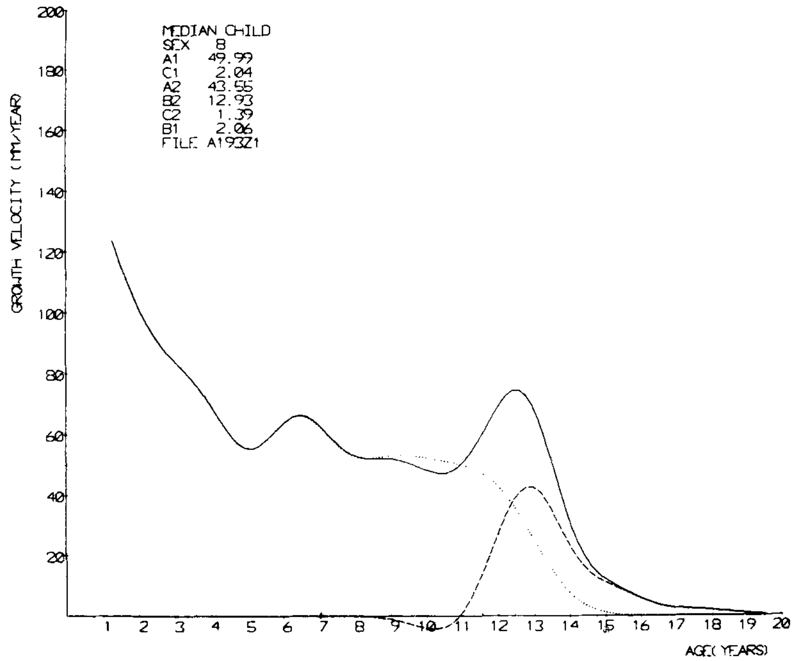


Figure 10. Growth velocity curve for median girl, i.e. with the median of the parameters inserted into the switch-off SIM.

means and standard deviations of individual parameters (table not given). In what follows we used shape functions based on the total sample of $N=90$ children (which is practically identical to the average of the estimates for three groups of $N=30$).

Since shape functions are not easy to interpret, we transformed them into a 'median child', by inserting for the individual parameters of (5) the respective median, for boys and girls separately. The result is shown in figures 9 and 10. There is a clear midspurt wholly attributed to the first component, occurring earlier in girls than in boys. Whereas quantitative features of the midspurt may become enhanced when improving the model, there is no doubt about the existence of the midspurt (modelling with fewer knots brought it out, too; compare also figure 8). Let us note that the difference in age between the midspurt peak and the pubertal peak is exactly the same for both sexes. For a discussion of the midspurt for other variables than height, consult Molinari *et al.* (1980).

Also remarkable is a dip in the sum of the two components before puberty starts. The second component, consisting of the pubertal peak, is truly dissociated from pubertal growth; it is not symmetric but quite heavily skewed towards higher ages.

(*d*) *Preliminary analysis of the individual parameters*

We first look at the parameters of both the additive (4) and the switch-off (5) model for the first random sample with $N=30$. Table 6 contains means and standard deviations for both sexes. The most striking result is that for the switch-off model, only the parameters of the second component are significantly different for boys and girls, whereas for the additive model the parameters of the first component ('prepubertal parameters') become significantly different, too.

Table 6. Switch-off versus additivity: a comparison of SIM parameters for boys ($N=15$) and girls ($N=15$) (means and (below) standard deviations).

Model		a_1	b_1	c_1	a_2	b_2	c_2
Switch-off	♂	45·2	2·83	3·67	68·7	14·2	1·34
		7·0	1·05	2·17	11·1	0·82	0·16
	♀	47·0	2·79	2·56	46·9	13·1	1·27
		7·4	1·41	1·15	9·5	0·73	0·30
Additive	♂	923	−7·29	5·45	280	13·3	0·99
		404	3·84	0·62	46	0·72	0·19
	♀	941	−5·19	4·54	221	11·7	1·22
		418	4·72	1·09	31	1·0	0·67

MANOVA for equality of a_1, b_1, c_1, for both sexes: $P=0{\cdot}27$ for switch-off, $P=0{\cdot}034$ for additive model.

MANOVA for equality of a_2, b_2, c_2 for both sexes: $P<10^{-5}$ for switch-off and additive model.

Table 7 does the same for the switch-off model and the $N=90$ sample. Tested in a univariate way, the prepubertal parameters a_1, b_1, c_1 are significantly different. This does not apply to multivariate testing (Wilks). The difference can probably be attributed to the different timing of the midspurt. It is notable that a_2 and b_2, related to the intensity and to the timing of the pubertal spurt, are strikingly different for boys and girls, whereas c_2, related to the duration of the spurt, was not significantly different. The rather large variance of b_1, c_1 may result from the fact that the two parameters separately are not well determined, as discussed below.

Table 7. SIM-parameters for boys and girls, switch-off model.

	Boys ($N=45$)		Girls ($N=44$)		
Parameter	Mean	Std. Dev.	Mean	Std. Dev.	t-statistic
a_1	46.0	6.9	49.8	7.1	−2.57*
b_1	2.81	1.08	2.26	0.98	2.53*
c_1	3.03	1.62	2.34	1.17	2.30*
a_2	66.4	10.2	46.4	12.2	8.45***
b_2	14.4	0.94	12.9	0.83	7.94***
c_2	1.35	0.20	1.41	0.28	−1.17

MANOVA (a_1, b_1, c_1): $P=0.0825$.
MANOVA (a_2, b_2, c_2): $P<10^{-5}$.

Measured by skewness and kurtosis, the univariate normality of b_1, c_1 is very much in doubt (table 8). The best explanation we can offer is that the two parameters are rather ill-determined: compare the start model (7), where only two out of the three parameters a_1,b_1,c_1 in S_1 are well determined.

Table 8. Skewness and kurtosis for SIM parameters.

	Boys ($N=45$)		Girls ($N=44$)	
Parameter	skewness	kurtosis	skewness	kurtosis
a_1	0.02	−1.07*	−0.23	−0.13
b_1	0.65	−0.83	1.39**	2.07**
c_1	1.20**	1.30**	1.98**	0.49
a_2	−0.46	0.36	0.89*	1.40**
b_2	0.33	−0.59	−0.27	−0.48
c_2	0.08	−0.67	0.17	−0.37

Table 9. Correlations between SIM-parameters and adult height for boys ($N=45$) and girls ($N=44$).

	a_1	b_1	c_1	a_2	b_2	Boys c_2	H_a
a_1		−0.84	−0.77	−0.10	−0.10	−0.20	0.23
b_1	−0.86		0.67	−0.04	−0.03	0.12	0.14
c_1	−0.77	0.73		0.10	−0.15	0.19	0.00
a_2	−0.27	0.25	0.17		−0.39	−0.05	−0.06
b_2	−0.19	0.06	0.02	−0.34		−0.11	0.04
c_2	−0.19	0.09	0.16	−0.15	−0.09		0.07
H_a	0.10	0.14	0.06	0.27	0.27	−0.16	

Girls

$r(\alpha=0.01)$: $\pm 0.24, \pm 0.25$.

The high correlations among the prepubertal parameters a_1, b_1, c_1 (table 9) may be somewhat artificial for the ill-determinacy outlined above. There is little correlation between prepubertal and pubertal parameters. As to the pubertal parameters, the distinct negative correlation between a_2 and b_2 is remarkable; it indicates that late-maturing children have a tendency to have a smaller pubertal peak superposed on previous growth. The single correlations on adult height (H_a) ar low, but the multiple correlation coefficients given in the following table may be more revealing:

	Boys	Girls
H_a on a_1,b_1,c_1	0·71	0·47
H_a on a_2,b_2,c_2	0·09	0·48
H_a on $a_1,b_1,c_1,a_2,b_2,c_2,a_2.c_2$	0·86	0·84

4. Discussion

We attacked the problem of modelling human height growth from a different angle than previously used: instead of postulating some functional shape for the whole or part of the developmental span, we determined the morphology of height growth from a sample of curves. As described in section 2, random subsamples were formed to check the reliability of results obtained. The structure assumed is that of a two-component shape-invariant model (see equations (4) and (5)). Based on one subsample with $N=30$, a switch-off model proved to be more consistent with the data than an additive model; in the second model the velocities of the two components superpose, while in the first, prepubertal growth is smoothly terminated with the appearance of the pubertal spurt. Such a model is also in good accordance with the coupling of the closure of the epiphyses to puberty and with the phenomenology of some disturbances of growth, as for example:

(*a*) Eunuchoidism, prepubertal castration.
The pubertal component is missing, and non-pubertal growth continues till long after the age of 20. Full ossification of the bones may never be reached. Although the individuals lack pubertal growth spurt, they grow on the average as tall as normal individuals.

(*b*) Precocious puberty.
Without going into a discussion of the aetiology of precocious puberty, we may note that these patients are taller than normal during their precocious puberty because of the pubertal growth spurt. Their final height is smaller than normal since non-pubertal growth has been stopped early and, therefore, contributes a reduced amount to final height.

The large overlap of non-pubertal and pubertal growth for the additive model (examples are given in figures 4–7) cannot be related to present endocrinological knowledge; sex steroids are present before puberty, although their role in general body growth before puberty is unknown. The analysis of individual parameters (tables 6 and 9) is also in favour of the switch-off model.

The decomposition of one observed series into two unobserved components cannot be done without some arbitrariness (this point has also been made by Tanner *et al.* (1976), when fitting the logistic curve for adolescence only). We do not claim to have an entirely correct identification of prepubertal and pubertal height velocity; the negative velocities appearing in the second component (figure 8) are a sign of inadequacy. We attribute this to our choice of a rigid coupling between the two components via the function φ.

The conclusion that the same model structure is appropriate for boys and girls can be drawn without ambiguity. This is rather surprising given the different endocrinological steering mechanisms in puberty, and given the different residual mean squares consistently found for boys and girls.

The pronounced midspurt may be the most striking feature of the growth pattern determined by SIM (figures 8–10); whereas its quantitative significance is low, it poses intriguing questions to endocrinologists. The previously applied smoothing approach (Largo *et al.* 1978) led to the postulate of the midspurt (confirmed by Molinari *et al.* 1980), and of a dip before the onset of the pubertal peak, now verified by SIM; in Tanner *et al.* (1966) such a dip was considered controversial, associated with late maturers, and with boys rather than with girls (compare also Bayley 1956). In the SIM it was present for both sexes, but more pronounced for boys, and there was no relation to late maturation (as judged from the graphic display). Yet early maturers seem to have a smoother growth with less random fluctuations. It remains to be seen if the finding of an exact equality for boys and girls of age difference between the midspurt and the pubertal spurt has a fruitful interpretation. It is an asset of the SIM approach that a midspurt could be taken into account within a six-parameter model (in the discussion of Preece (1978), who did not model the midspurt, an increase by two parameters was considered necessary for the midspurt; the additional 14 spline parameters to be estimated for the shape functions are negligible compared to the $6N = 540$ individual parameters, and if distributed evenly on individuals would lead to less than 6·6 parameters per individual).

The analysis of the individual parameters was preliminary, and the interpretation of results related to the prepubertal parameters a_1, b_1, c_1 has to be cautious due to ill-determinacy. The set of pubertal parameters is different for boys and girls in a highly significant way, but not the set of prepubertal parameters.

The parameter b_2 is related to age of peak height velocity (APHV), but since PHV relates to the superposition of both components, APHV has to be lower than b_2. The difference in b_2 for boys and girls is 1·5 years, but varies from 1·7 to 2·3 years for APHV in various investigations (table 1 of Preece 1978). The parameter a_2—a measure of the intensity of the spurt—roughly corresponds to PH defined in Largo *et al.* (1978) but its value is intermediate between PH and PHV (male: PH = 4·8, PHV = 9·0, $a_2 = 6{\cdot}6$; female: PH = 2·4, PHV = 7·1, $a_2 = 4{\cdot}6$). The parameter c_2, related to the duration of the spurt, is comparable to the parameter b of Tanner *et al.* (1976); in their and in our investigation, the difference for boys and girls was not significant. The descriptive parameter 'peak-basis' (PB) of Largo *et al.* (1978), which is not directly comparable to b_2, was significantly different for the two sexes.

Notable regarding the correlations between parameters is the approximate independence of pubertal parameters from prepubertal parameters. For pubertal parameters, the correlations between a_2 and b_2 are the only ones which were significant (male: −0·39; female: −0·34). This is in line with related but differently defined parameters: Tanner *et al.* (1976) found male: −0·50; female: −0·29 (PV, age at PV, table 4), Largo *et al.* (1978), for PHV, APHV male: 0·36; female: 0·36. Of great interest, but difficult to explain, are the differences in multiple correlations on adult height for boys and girls: whereas adult height was about equally determined by prepubertal and pubertal parameters for girls, only prepubertal parameters were good predictors for boys. Until further elucidation, a statistical artefact should not be excluded. Before individual parameters can be further analysed, it seems necessary to attack the problem of ill-determinancy of the three prepubertal parameters and probably reduce them to two. The remaining five parameters are expected to be interpretative to a large extent, and the derivation and analysis of secondary parameters (as defined in Largo *et al.* 1978) could confirm this. Our feeling is that where the assignment of interpretative parameters to individual curves and not modelling is the goal. the rather simple smoothing approach (Largo *et al.* 1978) may be of about equal merits compared to SIM.

Appendix

(i) *Distances and velocities*

We assume that the true residuals ε_{ij} of distances for individual i at age t_j have the following variance and correlation structure:

$$\mathrm{Var}(\varepsilon_{ij}) = \sigma_{ij}^2 \tag{1}$$

$$\mathrm{Corr}(\varepsilon_{ij}, \varepsilon_{ik}) = 0, \quad j \neq k. \tag{2}$$

For the residuals e_{ij} of velocities at age $(t_j + t_{j+1})/2$ we obtain

$$\mathrm{Var}(e_{ij}) = (\sigma_{ij}^2 + \sigma_{ij+1}^2)/(t_{j+1} - t_j)^2 \tag{3}$$

$$\mathrm{Corr}(e_{ij}, e_{ik}) = \begin{cases} 0, & k \neq j-1, j+1 \\ -\dfrac{\sigma_{ij+1}^2}{[(\sigma_{ij}^2 + \sigma_{ij+1}^2)(\sigma_{ij+1}^2 + \sigma_{ij+2}^2)]^{1/2}}, & k = j-1, j+1 \end{cases} \tag{4}$$

If the individual variability does not alter drastically with age, we have Corr $(e_{ij}, e_{ij+1}) \approx -\frac{1}{2}$ in good approximation, and the equality is exact if homoscedasticity within individuals (σ_{ij}^2 independent of j) holds. The computer program has an option to take into account this correlation structure (generalized least squares). No use was made of this option in the final computations, because preliminary tests showed no difference in the global results (RMS, shape functions) but some strange behaviour in individual cases. A possible reason for this is outlined in Stützle (1977; 3.10). The variance structure was in all cases taken into account (weighted least squares).

To prevent the mixing of aligned residuals with *a priori* different variances, residuals of half-year velocities were averaged for the purpose of aligned residual plots and age-specific Wilcoxon tests (figures 1–3; tables 1–3).

When fitting distances, the residual mean squares of individual i yields an estimate for σ_i^2 if variance is constant over age, and otherwise for

$$\frac{1}{n} \sum_{j=1}^{n} \sigma_{ij}^2.$$

If the model structure were not chosen correctly, the sum of pointwise squared bias terms would add. Following (3) we multiplied the residuals of velocities by $(t_{j+1} - t_j)/\sqrt{2}$ and this leads to an estimate of $(\frac{1}{2}\sigma_{i1}^2 + \sigma_{i2}^2 \ldots + \sigma_{in-1}^2 + \frac{1}{2}\sigma_{in}^2)/(n-1)$, if heteroscedastic, or σ_i^2, if homoscedastic. As long as the bias is negligible, our values of RMS are therefore comparable to those of distances; for an incorrect model as the double logistic, this is not the case, since the bias term for distances and velocities is not comparable; the double logistic model for velocities (table 4) gave RMS of 38 mm^2 (boys) and 37 mm^2 (girls), whereas the distance fit gave RMS of 122 mm^2 (boys) and 94 mm^2 (girls), in line with the results of Preece and Baines (1978) and Bock *et al.* (1973).

(ii) *Function spaces for SIM*

When improving an *a priori* fixed two-component SIM ('start model') S_1^0; S_2^0

we choose the corrections γ_1, γ_2 from vector spaces Γ_1, Γ_2:

$$S_1(x)=S_1^0(x)+\gamma_1(x)$$
$$S_2(y)=S_2^0(y)+\gamma_2(y) \tag{5}$$

(x,y: standardized age scales for 1st, 2nd component)

The classes Γ_1, Γ_2 chosen were cubic spline functions with knots $x_1,\ldots,x_k$ or $y_1,\ldots,y_m$, vanishing outside $[x_1,x_k]$ or $[y_1,y_m]$.

Splines (rather than ordinary polynomials) were chosen since their behaviour is determined locally, but with a smooth transition, and because of good numerical properties. Each element of $\Gamma_1[(k-4)$-dimensional] and $\Gamma_2[(m-4)$-dimensional] has a unique representation in terms of B-splines (de Boor 1978):

$$\gamma_1(x)=\sum_{j=1}^{k-4}\alpha_j B_j(x;k)$$
$$\gamma_2(y)=\sum_{j=1}^{m-4}\beta_j B_j(y;m) \tag{6}$$

The pair of coordinate vectors (α, β) defines the correction functions γ_1, γ_2 (in contrast to the individual parameters $a_1, \ldots c_2$ these population parameters do not increase in number with sample size). We choose 10 knots for the first and 12 knots for the second component (equidistant for computational convenience):

$$\begin{array}{lll}\Gamma_1{:}x_1=-11{\cdot}5, & x_i=x_{i-1}+0{\cdot}65 & (i=2,\ldots,10)\\ \quad[-0{\cdot}6y] & \quad[1{\cdot}7y] & \\ \Gamma_2{:}y_1=-5{\cdot}5, & y_i=y_{i-1}+1{\cdot}0 & (i=2,\ldots,12)\\ \quad[6{\cdot}1y] & \quad[1{\cdot}4y] & \end{array} \tag{7}$$

In square brackets: years for the average child, i.e. the child with average parameters corresponding to absolute ages x,y.

Some arbitrariness lies in the choice of knots; to bring out fine structure possibly present in the data, we took rather too many than too few. To check for overfitting, we subdivided the individuals into three random samples (figure 8) and also redid the computations with fewer knots (7 and 9). There were no substantial changes (the midspurt peak became reduced but did not disappear, and this reduction also led to a reduction in the differences for boys and girls for the prepubertal parameters; individual curves looked smoother and more stable).

(iii) *The algorithm*

The following iterative procedure was used to solve the two interdependent problems of estimating the shape functions S_1, S_2 and the individual parameters $\{a_{1i},\ldots,c_{2i}: i=1,\ldots,N\}$:

First step:

A start model with shape functions S_1^0, S_2^0 was chosen (equation (7) or (8)). We then estimated individual parameters $a_1^{(0)},\ldots,c_2^{(0)}$ with respect to S_1^0, S_2^0 by non-linear regression (program NL2SOL; Dennis *et al.* 1978). When convergence as defined by the program was reached, 4–5 decimals of the solution and the RMS did not change.

pth step ($p \geqslant 2$):

Now the best fitting spline function (following (i)) to the pooled residuals of all N individuals was determined by linear regression:

$$\sum_{i=1}^{N} \sum_{j=1}^{n-1} [v_i(t_j^*) - a_{1i}^{(p-2)} \sum_{S=1}^{k} \alpha_S^{(p-1)} B_S(x_{ij}^{(p-2)}) - a_{2i}^{(p-2)} \sum_{S=1}^{m} \beta_S^{(p-1)} B_S(y_{ij}^{(p-2)})]^2 = \min$$

$$x_{ij}^{(p-2)} = \frac{t_j^* - b_{1i}^{(p-2)}}{c_{1i}^{(p-2)}}, \quad y_{ij}^{(p-2)} = \frac{t_j^* - b_i^{(p-2)}}{c_{2i}^{(p-2)}}$$

The determination of the 14 variables ($\alpha^{(p-1)}, \beta^{(p-1)}$) characterizing S_1^{p-1}, S_2^{p-1} for up to 2700 observations ($N=90, n=30$) is a formidable task: the solution by sequential Householder transformation (Lawson and Hanson 1974) is numerically sound and only (NVAR+2) rows of the design matrix need to be in core (NVAR = number of variables).

Then non-linear regression with respect to S_1^{p-1}, S_2^{p-1} was performed yielding individual parameters $a_1^{(p-1)}, \ldots, c_2^{(p-1)}$.

The iteration was terminated when the change in all of the parameters ($\alpha^{(p-1)}$, $\beta^{(p-1)}$) was less than their standard deviation, which usually resulted in five iterations. Judged from the change in RMS, three steps would probably suffice.

The statistical stability of (α, β) depends on N, and there seemed to be a major improvement when going from $N=15$ to $N=30$.

Acknowledgement

Dr Stützle was supported partly by the Fritz Hoffmann La Roche Foundation.

References

Bayley, N., 1956, Growth curves of height and weight by age for boys and girls, scaled according to physical maturity. *Journal of Pediatrics*, **48**, 187–194.

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

De Boor, C., 1978, *A Practical Guide to Splines* (Berlin: Springer-Verlag).

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.

Dennis, J. E., Gay, D. M., and Welsch, R. E., 1978, An adaptive nonlinear least-squares algorithm. Technical Report 142, M.I.T. Operations Research Center, Cambridge, Ma.

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

Falkner, F., 1960, *Child Development. An International Method of Study*. Modern Problems in Pediatrics, (Basel: Karger).

Largo, R. H., Stützle, W., Gasser, Th., Huber, P. J., and Prader, A., 1978, Analysis of the adolescent growth spurt using smoothing spline functions. *Annals of Human Biology*, **5**, 421–434.

Lawson, C. L., and Hanson, R. J., 1974, *Solving Least Squares Problems* (Hemel Hempstead: Prentice Hall).

Lawton, W. H., Sylvestre, E. A., and Maggio, M. S., 1972, Selfmodeling nonlinear regression. *Technometrics*, **14**, 513–532.

Marubini, E., Resele, L. F., and Barghini, G., 1971, A comparative fitting of the Gompertz and logistic functions to longitudinal height data during adolescence in girls. *Human Biology*, **43**, 237–252.

Molinari, L., Largo, R., and Prader, A., 1980, Analysis of the growth spurt at age 7 (mid-growth spurt). *Helvetica Paediatrica Acta*, **35**, 235–334.

Preece, M. A., 1978, Analysis of the human growth curve. In *Pediatrics and Growth*, edited by D. Barltrop (London: Fellowship of Postgraduate Medicine), pp. 77–90.

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

Smith, D. W., Truog, W., Rogers, J. E., Greitzer, L. J., Skinner, A. L., McCann, J. J., Sedgwick Harvey, M. A., 1976, Shifting linear growth during infancy: illustrate of genetic factors in growth from fetal life through infancy. *Journal of Pediatrics*, **89**, 225–230.

STÜTZLE, W., 1977, Estimation and parameterization of growth curves. Thesis No. 6039, ETH, Zürich.

TANNER, J. M., WHITEHOUSE, R. H., and TAKAISHI, M., 1966, Standards from birth to maturity for height, weight, height velocity and weight velocity: British children 1965. *Archives of Disease in Childhood*, **41**, 454–471; 613–635.

TANNER, J. M., WHITEHOUSE, R. H., MARUBINI, E., and RESELE, L., (1976), The adolescent growth spurt of boys and girls of the Harpenden Growth Study. *Annals of Human Biology*, **3**, 109–126.

Address correspondence to: Dr Th. Gasser, Zentralinstitut für Seelische Gesundheit, P.O. Box 5970, 6800 Mannheim 1, FR Germany.

Zusammenfassung. Wir stellen einen neuen Ansatz zur Modellbildung des Höhenwachstums des Menschen vor, der auch für andere Variablen geeignet ist. In einem mathematischen Algorithmus wird eine *a priori* Vorstellung über die funktionelle Form dieser Entwicklung unter Benützung der Daten verbessert; das resultierende 2-Komponenten 'shape-invariant model' (SIM) erlaubt eine approximativ biasfreie Anpassung an longitudinale Daten von 1 bis 20 Jahre, wobei jeder individuelle Verlauf durch 6 Parameter charakterisiert ist. Der SIM-Ansatz erwies sich als geeignet, grundlegende qualitative Fragen zu beantworten. Bei einem Vergleich zwischen einem 2-Komponenten SIM, das additiv ist, und einem, bei dem das Erscheinen der Pubertät weiteres Wachstum der nicht-pubertären Komponente inhibiert ('Abschalt-Modell'), erwies sich letzteres in verschiedener Hinsicht als überlegen. Die Modellstruktur, sowohl vor als auch während der Pubertät, wurde für Jungen und Mädchen als identisch geschätzt. Erwähnenswert an qualitativen Merkmalen der gefundenen Modellstruktur ist: ein ausgeprägter Zwischenspurt, ein Absinken der Geschwindigkeit vor dem Pubertätsschub and eine deutliche Asymmetrie des Pubertätsgipfels. Eine vorläufige statistische Analyse der individuellen Parameter bestätigt einige mit anderen Methoden erzielte Ergebnisse bezüglich Geschlechtsunterschiede und Assoziationen zwischen Parametern.

Résumé. Nous présentons une nouvelle méthode pour l'estimation des courbes de croissance en taille des garçons et des filles, qui pourrait être aussi appliquée à d'autres paramètres. Une estime initiale de la forme fonctionelle de ces courbes est améliorée d'une façon consistante avec les données par une procédure mathématique récursive. Le modèle qui en résulte, appelé shape-invariant (SIM) à deux composants permet une description dépourvue de bias des courbes de croissance de 1 à 20 ans, chaque courbe individuelle étant caractérisée par 6 paramètres. Le modèle shape-invariant est de grande utilité dans une discussion qualitative de modèles biomathématiques de la croissance. Aussi l'application du modèle aux courbes de vitesse, plutôt qu'aux distances donne une meilleure idée de ces phénomènes. Nous comparons deux modèles SIM à deux composants, le premier est additif, dans le deuxième (switch-off SIM) l'apparition du composant 'pubertaire' cause une réduction et, après un certain délai, la disparition du composant non pubertaire. Le modèle switch-off paraît supérieur de bien des points de vue au modèle purement additif et est discuté dans la suite du travail. Il se trouve que le même modèle (shape functions) s'applique chez les garçons et chez les filles, aussi bien avant que pendant la puberté. Les autres aspects intéressants du modèle sont: une poussée de croissance prépubertaire (midspurt growth) une anfractuosité avant le début de la poussée de croissance pubertaire, une très claire asymmétrie au pic correspondant. Une analyse statistique préliminaire des paramètres individuels confirme les résultats de travaux précédants à propos des différences entre les sexes et des relations de ces paramètres entre eux et avec la taille adulte.