ANNALS OF HUMAN BIOLOGY, 1984, VOL. 11, NO. 5, 397–411

Velocity and acceleration of height growth using kernel estimation

T. GASSER

Zentralinstitut für Seelische Gesundheit, Mannheim

and W. KÖHLER, H-G. MÜLLER and A. KNEIP

Institut für Angewandte Mathematik, Universität Heidelberg

and R. LARGO, L. MOLINARI and A. PRADER

Kinderspital, Universität Zürich

Received 28 February 1983; revised 26 March 1984

Summary. A method is introduced for estimating acceleration, velocity and distance of longitudinal growth curves and it is illustrated by analysing human height growth. This approach, called kernel estimation, belongs to the class of smoothing methods and does not assume an a priori fixed functional model, and not even that one and the same model is applicable for all children. The examples presented show that acceleration curves might allow a better quantification of the mid-growth spurt (MS) and a more differentiated analysis of the pubertal spurt (PS). Accelerations are prone to follow random variations present in the data, and parameters defined in terms of acceleration are, therefore, validated by a comparison with parameters defined in terms of velocity. Our non-parametric-curve-fitting approach is also compared with parametric fitting via a model suggested by Preece and Baines (1978).

1. Introduction

A new method is presented for estimating distance, velocity and acceleration of longitudinal growth data, such as height. Given the variability present (figures 3–6), the accurate determination of acceleration presents some difficulties; the acceleration curve could, however, be of interest in addition to distance and velocity. Aspects of regulation of growth are better accessible via accelerations (Gasser, Köhler, Müller, Largo, Molinari and Prader 1983 b), and some features like the mid-growth spurt (MS) are easier to quantify (see below and Gasser, Müller, Köhler, Prader, Largo and Molinari 1983 a). The method proposed, which is well established in mathematical statistics, is called kernel estimation and consists of applying weights (differently chosen for distance, velocity and acceleration) to the measurements, and averaging the weighted measurements over appropriate age intervals (more details are given in *Subjects and methods*).

This method is non-parametric as was the heuristic graphical procedure by Tanner, Whitehouse and Takaishi (1966) and the spline smoothing approach by Largo, Gasser, Prader, Stützle and Huber (1978). It is called non-parametric, since no fixed parametric function, such as the logistic one, needs to be postulated for modelling growth (Marubini 1978, Preece and Baines 1978, Stützle, Gasser, Molinari, Prader and Huber 1980). Whenever an (unknown) parametric model is mis-specified, its fitting to the data will reflect this mis-specification; the resulting errors will be inflated when velocity is obtained by differentiating a fitted-distance curve, as shown by El Lozy (1978). For this reason, a non-parametric method was preferred to a parametric one. Cubic smoothing splines (as used by Largo *et al.* 1978) are closely related to kernel estimates. One disadvantage is that their accelerations are piecewise linear (the same holds for variable

knot cubic splines as used by Berkey, Reed and Valadian 1983). They are also computationally more expensive and less easily amenable to modifications for specific purposes. In Gasser, Müller, Köhler, Molinari and Prader (1984), it has also been shown that they are statistically somewhat inferior in terms of mean squared error for analysing growth curves.

The models proposed by Preece and Baines (1978) proved to be a major step forward for analysing growth, and, being the best parametric models known so far, they are fair candidates for a comparison with kernel estimates. Classical parameters which characterize the pubertal spurt (PS) will be determined from the curve fitted using model 3 of Preece and Baines (1978), and also from the velocity and the acceleration curves obtained by kernel estimates. Further steps of validation of the new method are a comparison with cross-sectionally averaged increments and a discussion of several examples.

The quantification of the MS, taking place at about seven years of age, was an important goal when starting the methodological work. The MS was verified statistically (but not quantified) by Molinari, Largo and Prader (1980) and, with some interesting differences, by Tanner and Cameron (1980). We expected that it would be easier to quantify the MS from acceleration curves: in the velocity curve the MS is usually a small peak riding on the declining trend of pre-pubertal velocity. In many instances there is no true maximum in velocity at the MS peak (note that such a maximum is used when identifying a spurt from the velocity curve). The acceleration curve, however, might show clear-cut peaks; it is not decisively affected by the declining velocity trend: if the trend is linear around seven years (compare figures 4 and 5 of Berkey *et al.* 1983), the accleration of the MS is just shifted by a negative constant, and otherwise keeps its pattern.

2. Subjects and methods

In 1954 a prospective longitudinal study of growth and development of 414 healthy Swiss children was initiated at the Kinderspital, Zürich, in an internationally coordinated design (Falkner 1960). For the present work, children are excluded if height measurements are incomplete from birth to three years, or if later two consecutive or more than three observations are missing. None of the children included had a disease hampering growth. A random sample of 45 boys and 45 girls was drawn from those children fulfilling these criteria.

Measurements

Children were measured at 4, 13, 26 and 39 weeks, then at 1, 1·5 and 2 years, and after the age of 2 years annually at birthdays ± 14 days up to 9 years for girls and up to 10 years for boys. Afterwards they were measured every six months within the same limits until the annual increment in height was less than 0·5 cm per year. Yearly measurements were continued until the increment had become less than 0·5 cm in two years.

Editing

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

Computations

Computing was done on the IBM 370/168 of the Computer Center of the University of Heidelberg. The programs for kernel estimation, for parametrization, and for graphical analysis are our development. Non-linear least squares for fitting model 3 of Preece and Baines (1978) was based on an algorithm by Deuflhard and Apostolescu (1980). Further processing of the parameters was performed mainly via the program package SAS.

Kernel estimation

When height H is measured at ages $t_1, \ldots t_N$ (for child no. j) the following regression model is assumed:

$$H_j(t_i) = H_j^*(t_i) + \epsilon_j(t_i) \qquad \begin{array}{l} i = 1, \ldots, N \\ j = 1, \ldots, n \end{array}$$

where:

$H_j(t_i)$ = height measured at age t_i of subject j
$\epsilon_j(t_i)$ = random variation of height measurements
$H_j^*(t_i)$ = true height at age t_i

The distance curve $H^*(t)$, the velocity $\mathrm{d}H^*/\mathrm{d}t$, and the acceleration $\mathrm{d}^2H^*/\mathrm{d}t^2$ (the first and second derivative of distance, respectively) are of interest. The random variations $\epsilon(t)$ are due to measurement error, seasonal and diurnal change, environmental conditions, etc. A crude estimate of the standard deviation of ϵ was obtained by local polynomial fitting and by cross-validation: both gave estimates of approximately 5 mm from 2–8 years and approximately 3 mm from 15–18 years. The parametric model fitted to the data by the principle of least squares is of the following form (Preece and Baines 1978, model 3):

$$H_{\mathrm{PB}}(t) = a - \frac{4(a - H(b))}{(\exp(c(t-b) + \exp(\mathrm{d}(t-b)))\ (1 + \exp(e(t-b)))}$$

Kernel estimation was introduced into mathematical statistics by Rosenblatt (1956) for probability densities, and later extended to regression functions (for a review of the statistical literature, see Collomb 1981). A kernel estimate for H^* was defined and studied in Gasser and Müller (1979) and later generalized for determining derivatives from noisy data (large sample results are given in Gasser and Müller (1984); the choice of kernels is discussed in Gasser, Müller and Mammitzsch (1985); data-analytic problems with respect to growth curves have been dealt with in Gasser *et al.* (1984). In contrast to parametric regression, it does not assume a unique pattern of growth for all subjects and treats the development of each child in his own right. The estimate $\hat{H}_\nu$ for determining $\mathrm{d}^\nu H^*/\mathrm{d}t^\nu$ ($\nu = 0,1,2$) is of the following form:

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

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

Here ν = 0,1,2 (distance, velocity, acceleration)
W_ν = kernel ('weighting function') appropriate for ν-th derivative
b = bandwidth (smoothing parameter)
s_t = $(t_{i+1}+t_i)/2$
t = age at which $d^\nu H^*/dt^\nu$ is to be determined

Kernel estimation consists, therefore, in the weighted averaging (with weights g_i) of height measurements. These weights are non-zero for measurements $H(t_i)$ falling in the age interval $(t-b,t+b)$ since the kernal W_ν stretches from $(t-b)$ to $(t+b)$, with its centre at t. A weight g_i for the measurement at t_i is then computed from the kernel function W_ν (compare figure 1) by taking the area under the curve between two time points half-way to the adjacent measuring times (note that the weights can be negative). The integral of $W\nu$ can be obtained analytically which makes computing much easier. For $\nu=0$, this estimate is a plausible one, consisting in weighted moving averaging which leads to some smoothing of the raw height measurements. In order to obtain smooth derivatives from noisy data, specific kernels were developed from mathematical principles. These kernels have some intuitive appeal. The kernel for $\nu=1$ effectively averages central differences (with respect to t), rather than increments, and it puts higher weights with opposite signs on values lying somewhat distant from the age t (compare figure 1).

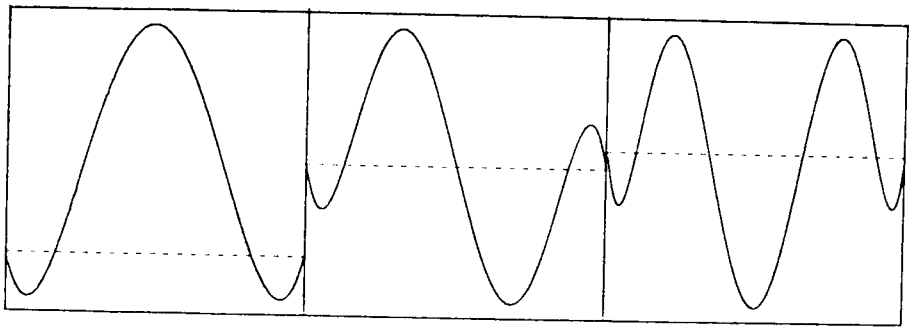

Figure 1. Kernels used for estimating the distance (left), velocity (middle) and acceleration curves (right). Every kernel on its own scale. Dotted line = zero.

In the definition of $\hat{H}_\nu$, the smoothing parameter b and the kernel function W_ν are at our disposal (restricted by some mathematical conditions in the latter case). Kernels found to be mathematically optimal were also superior (with respect to mean squared error) for analysing growth curves (Gasser *et al.* 1984); their analytical form for $\nu=0,1,2$ is the following (see also figure 1):

$$W_0(x) = \begin{cases} \frac{15}{32}\,(7x^4-10x^2+3) & |x| \leqq 1 \\ 0 & |x| > 1 \end{cases}$$

$$W_1(x) = \begin{cases} \frac{105}{32}\,(-9x^5+14x^3-5x) & |x| \leqq 1 \\ 0 & |x| > 1 \end{cases}$$

$$W_2(x) = \begin{cases} \frac{315}{64}\,(77x^6-135x^4+63x^2-5) & |x| \leqq 1 \\ 0 & |x| > 1 \end{cases}$$

A good choice of the smoothing parameter b, enabling a compromise between the smoothness of the curve obtained and the fidelity to the data, is crucial indeed. For $\nu = 0$ this choice might be guided by eye by comparing the kernel fit with the data. For $\nu = 1,2$ there are no measurements and the properties of increments (see next section) make such a subjective choice of b by eye even more dubious than for $\nu = 0$. The finite sample evaluation technique developed in Gasser *et al.* (1984) allowed us to minimize the integrated mean square error, pooled over subjects, which provided more rational choices for b (to the eye, they are rather conservative). For $\nu = 0/1/2$, these were $b = 3\cdot4/3\cdot8/4\cdot0$ years in the pre-adolescent and $b = 1\cdot8/3\cdot1/3\cdot6$ years in the adolescent period, with smooth transitions in between. These varying bandwidths resulted from our requirement of constant variability of $\hat{H}_\nu$ over age despite unequal measurement intervals (compare Gasser *et al.* 1984, section 6).

Properties of increments

Properly standardized first and second increments can serve as raw velocity RV and as raw acceleration RA at midpoints (Δt = measurement interval):

$$RA\left(\frac{t_i + t_i - 1}{2}\right) = (H(t_i) - H(t_{i-1}))/\Delta t$$

$$RV\ (t_i) = (H(t_{i+1}) - 2H(t_i) + H(t_{i-1}))/\Delta t^2$$

The second formula has to be modified when $(t_{i+1} - t_i)$ and $(t_i - t_{i-1})$ are not of equal size. Statistical properties of raw velocity and acceleration are not simple: assuming that the random term ϵ has variance σ^2, the variance of the raw velocity is $2\sigma^2/\Delta t^2$ and the correlation between consecutive ages is $-0\cdot5$ and zero otherwise. As a consequence, the variance of RV is quadrupled for half-year intervals compared to yearly ones. The variance of RA is $6\,\sigma^2/\Delta t^4$ and is, therefore, for half-year measurements 16 times the variance based on yearly measurements. Correlations for RA between neighbouring values are $-4/6$ and for values two apart $+1/6$. This demonstrates quantitatively that the random term becomes magnified when working with derivatives (compare also figures 3–6) which explains why the estimation of derivatives from non-exact data is a mathematical problem up to the present day.

Definition of parameters

The parameters used for comparing the different methods are the age of onset of the PS (called T6 here, consistent with our other work), the age of peak height velocity (T8) and the velocity V and height H reached at these ages. These parameters were extracted from the velocity curve obtained by kernels (denoted by T6(V), VT6(V) etc.), and from the velocity curve derived from the parametric fit (designated as T6(PB) etc.). This latter was achieved by computing the age of minimum velocity prior to the PS and the age of maximal velocity. The first point of zero acceleration (for the kernel estimate) provides T6(A), VT6(A), HT6(A), and the second one T8(A) etc.

Things are different with respect to the time of the end of the PS(T9), defined by Largo *et al.* (1978) as the point in the velocity curve where minimal pre-PS velocity was reached again. The acceleration curve leads to a more natural definition given by the age of maximal deceleration (= T9(A)). In spite of the difference in definition, a comparison of T9(V), T9(PB) and T9(A) is undertaken to check how far previous results are comparable.

Further parameters are shown in figures 2 and 9 but are not used otherwise here: T1 is age 4; T2, T3 and T4 characterize the individual age of maximal acceleration, maximal velocity, and maximal deceleration during the MS, and T5 designates the age where acceleration begins to rise before the PS.

3. Results

Velocities and accelerations

Figure 2 gives the velocity and the acceleration curve of a boy obtained by kernel estimates. He reached an adult height of 183·4 cm, his PS is peaking at 15·2 years (almost 1·5 years after the average boy) and he does not show any MS. A slightly asymmetrical velocity peak is sitting on a declining pre-PS velocity trend. Deceleration shows a sharp decline up to three years and afterwards a roughly constant one up to puberty; the acceleration pattern of the PS consists of a biphasis oscillation; deceleration is higher than maximal acceleration: here, and in the following examples, this difference shows lucidly the asymmetry of the velocity peak of the PS.

In the next example (figure 3) kernel estimates are compared with the parametric fit, and with raw velocity and acceleration (see methods) both showing an accentuated MS (girl with adult height = 167·2 cm). The low variability during the pre-PS period (yearly measurement intervals) facilitates a comparison between fitted curves and raw velocity and acceleration. The kernel estimate is quite satisfactory whereas the parametric fit does not reflect interesting facets of the data during the MS. The general pattern of the MS is quite similar to that of the PS, as can be inferred from the acceleration. The two methods agree quite well during the PS, apart from an earlier rising for the parametric fit.

A subjective judgment is more difficult in the next example (figure 4, a girl with adult height of 164·5 cm) due to a larger inherent variability of the data. Again, the main methodological difference lies in the ignoring of the MS by the parametric model

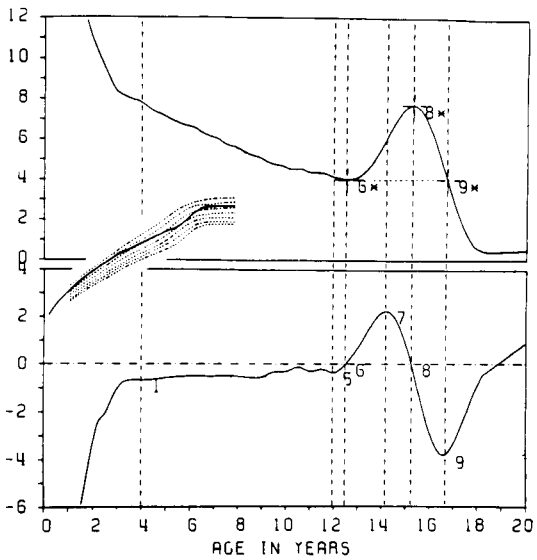

Figure 2. Velocity (above) and acceleration (below) of a boy. Dashed vertical lines = timings defined from the acceleration curve (given by T1, . . . ,T9; T6*, T8*, T9* are defined in the velocity curve). Small graph gives distance together with cross-sectional percentiles 3%, 10%, 25%, 50%, 75%, 90%, 97%.

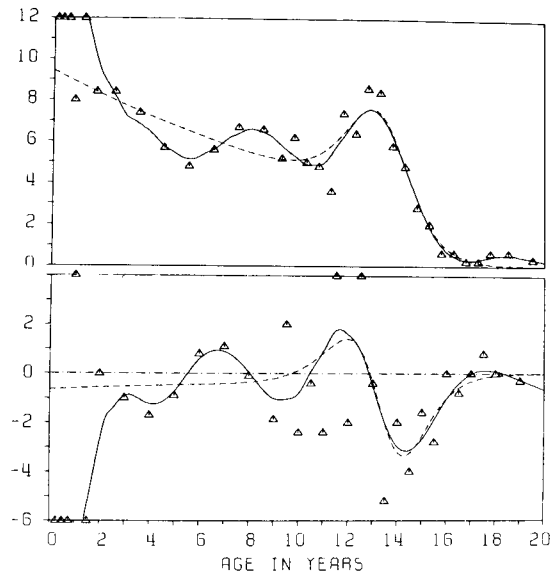

Figure 3. Velocity (above) and acceleration (below) of a girl. Solid line = kernel estimates, dashed line = parametric fitting. Triangles are raw velocities and acceleration (note higher variability for half-year intervals), if value is outside range, it is put at the delimiting rectangle.

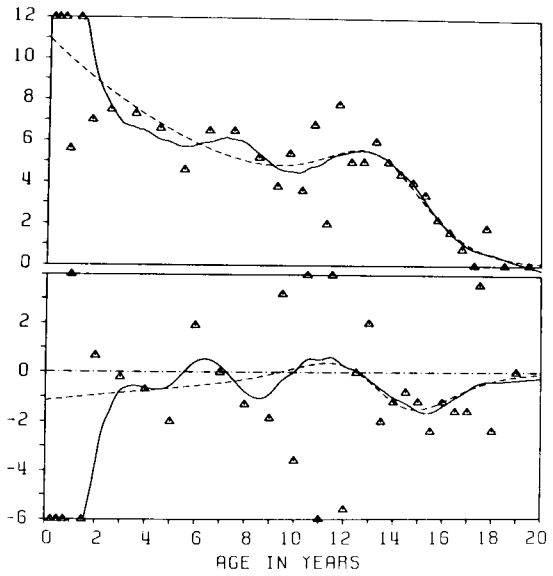

Figure 4. Legend as in figure 3, results for a different girl.

and its earlier rise to the PS. Note the large variability of raw accelerations, in particular for half-year measurements.

Figure 5, a boy with an adult height of 173·9 cm, shows a moderately large, rather broad MS in the 'data' and in the kernel estimate. Both methods show a tendency to underestimate the pubertal peak, the kernel estimate somewhat more than the parametric fit.

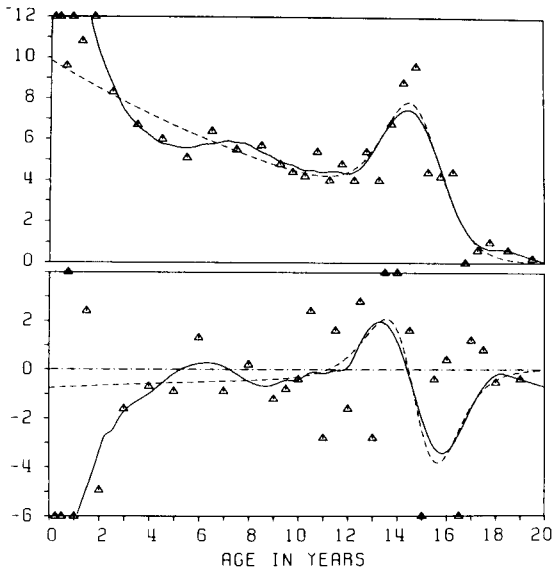

Figure 5. Legend as in figure 3, results for a boy.

The next example (figure 6, a girl) shows a disagreement between the parametric and the non-parametric method not only during the pre-PS period, but also for the PS. The high variability (in particular for acceleration) makes a subjective judgment difficult. This holds in particular for acceleration where discrepancies are most pronounced.

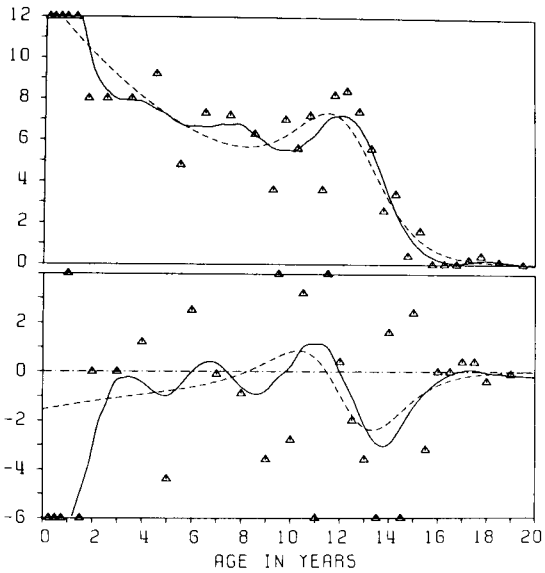

Figure 6. Legend as in figure 3, results for a boy.

Figure 7 (girls) and 8 (boys) give a comparison of raw velocities and acceleration with the respective kernel and parametric fits, averaged cross-sectionally over the sample of 45 boys and 45 girls (cross-sectional averaging was chosen since it does not rely on the

method investigated). The ups and downs in the raw velocities and raw accelerations reflect the correlation pattern of increments, as discussed in *Subjects and methods*. For either sex, a MS is clearly visible in the averaged raw 'data' and it is adequately described by kernel estimates. The visual discrimination of an MS is easier for boys due to a longer intervening period between the MS and the PS, whereas MS and PS follow closely for girls. The two methods show a good agreement with the average curves with respect to the PS with one exception: for girls, parametric fitting sets the onset of the PS

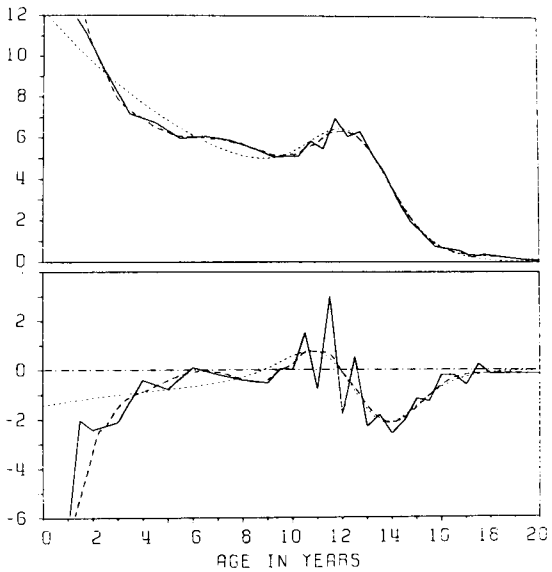

Figure 7. Velocity (above) and acceleration of $N = 45$ girls, averaged cross-sectionally. Solid line = raw velocity and acceleration ('data'), connected by straight lines between measurements. Dashed line = kernel estimates. Dotted lines = parametric fitting.

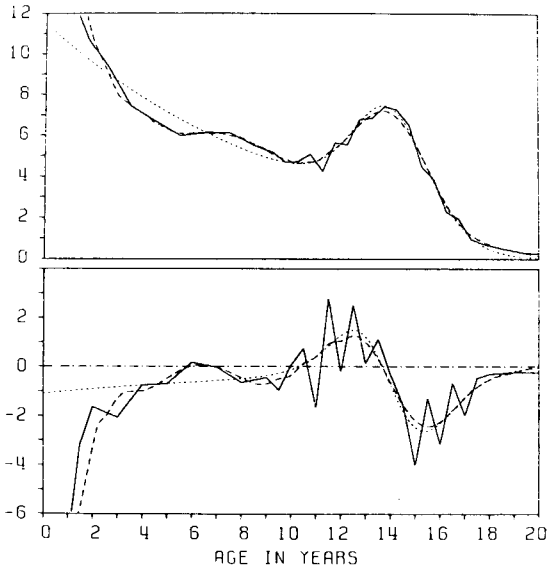

Figure 8. Legend as in figure 7, results for $N = 45$ boys.

at an earlier age than is to be expected from the data (to be seen most clearly as a displacement of the first zero of the acceleration curve).

The last example (figure 9, a rather small boy of 167·1 cm) shows two small bumps on the declining velocity trend at four and seven years which one would hesitate to designate as peaks. The acceleration curve, however, leads to an unequivocal quantification of the MS, and the previous figures give confidence that properties of the data are reflected in such a quantification.

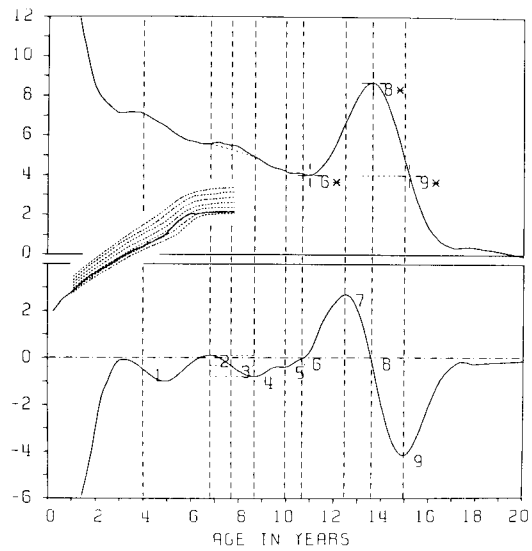

Figure 9. Legend as in figure 2.

A comparison of parameters of the PS

Classical parameters characterizing the PS were obtained by non-parametric acceleration (T6(A) etc.), by non-parametric velocity (T6(V) etc.) and via parametrically defined velocity (T6(PB) etc.) as outlined in *Subjects and methods*; tables 1 and 2 contain means and standard deviations, separately for boys and girls. Non-parametric velocity and acceleration curves lead to almost identical means and standard deviations with respect to the onset of the PS (T6, HT6, VT6) and its peaking (T8, HT8, VT8). Systematic differences arise when using parametric fitting (model 3 of Preece and Baines 1978): the age of onset of the PS is put at a younger age, in particular for girls, and, as a consequence, differences arise for velocity and height reached. The parameters of the peak itself show a closer agreement, except for the average velocity of boys which is lower by 3·8 mm/year for the kernel method. With regard to the parameters of the end of the PS, differences are somewhat larger, as to be expected (see *Subjects and methods*).

Rank correlations, used here as measures of agreement between different approaches to parametrization are given in table 3. Regarding the onset and the peak of the PS, the agreement of the parameters defined either from the non-parametric acceleration or velocity curve is excellent. When comparing the parametrization based on parametric fitting (model 3, Preece and Baines 1978) with the one based on non-parametric acceleration curves, the agreement is still quite good with respect to the peak

of the PS, but somewhat weaker for characterizing its onset. When keeping in mind that the end of the PS (T9 etc.) was previously defined rather arbitrarily (Largo *et al.* 1978) the agreement with the more natural definition in terms of acceleration is remarkably good, except for velocity at the end, where it is poor; this is due to the sharp decline of velocity at this time which can make drastic effects when age is varied.

Table 1. Comparison of methods for determining parameters of the adolescent spurt for $N=45$ boys.

Time	Quant.	TI(A)	TI(V)	TI(PB)	HTI(A)	HTI(V)	HTI(PB)	VTI(A)	VTI(V)	VTI(PB)
$I=6$	$\bar{x}$	10·90	10·85	10·62	143·6	143·4	142·1	4·33	4·31	4·48
	s	1·07	1·11	0·83	6·8	6·7	6·5	0·50	0·50	0·47
$I=8$	$\bar{x}$	13·91	13·92	14·00	161·4	161·7	162·6	8·31	8·32	8·70
	s	0·95	0·94	0·89	6·6	6·7	6·3	0·82	0·82	1·03
$I=9$	$\bar{x}$	15·36	15·55	15·47	172·4	173·1	172·8	4·99	4·31	4·48
	s	0·91	0·96	0·94	6·6	6·4	6·4	0·59	0·50	0·47

Table 2. Comparison of methods (A = non-parametric accelerations/V = non-parametric velocities/PB = velocity of fitting model 3 Preece and Baines) for determing parameters of the adolescent spurt for $N=45$ girls. T6/T8/T9 = onset/peaking/offset of PS.

Time	Quant.	TI(A)	TI(V)	TI(PB)	HTI(A)	HTI(V)	HTI(PB)	VTI(A)	VTI(V)	VTI(PB)
$I=6$	$\bar{x}$	9·73	9·74	9·01	136·3	136·3	132·3	4·85	4·83	4·90
	s	0·96	1·01	0·71	7·3	7·5	6·2	0·58	0·59	0·46
$I=8$	$\bar{x}$	12·22	12·22	12·09	150·4	150·6	150·2	7·00	7·00	7·09
	s	0·81	0·80	0·78	6·0	5·9	5·9	0·95	0·95	1·09
$I=9$	$\bar{x}$	13·84	13·56	13·42	160·6	159·3	158·5	4·13	4·83	4·90
	s	0·84	0·86	0·84	6·0	6·1	6·3	0·62	0·59	0·46

Table 3. Rank correlations as measures of concordance between different methods (non-parametric accelerations on one hand, non-parametric velocities and velocities following Preece and Baines (1978) on the other hand) of parametrization of the adolescent spurt ($N=45$ boys and $N=45$ girls). T6/T8/T9 = onset/peaking/offset of PS.

		TI(A)		HTI(A)		VTI(A)	
I	Parameters	M	F	M	F	M	F
6	T6,HT6,VT6(V)	0·95	0·99	0·99	0·99	1·00	1·00
	T6,HT6,VT6(PB)	0·79	0·81	0·82	0·89	0·82	0·82
8	T8,HT8,VT8(V)	1·00	1·00	1·00	0·99	1·00	1·00
	T8,HT8,VT8(PB)	0·97	0·87	0·94	0·89	0·91	0·97
9	T9,HT9,VT9(V)	0·96	0·95	0·97	0·96	0·18	0·32
	T9,HT9,VT9(PB)	0·95	0·83	0·95	0·89	0·29	0·22

4. Discussion

As shown by examples, accelerations bring out an interesting morphology of human height growth: in addition to the pubertal spurt (PS), a small spurt at about seven years old can be seen for most childlren, and can be detected and quantified more easily in the acceleration than in the velocity curve. The existence of such a mid-growth spurt (MS) had been previously verified as a systematic pattern of growth for various somatic quantities by Molinari *et al.* (1980) and by Tanner and Cameron (1980). Accelerations also shed new light on the PS: they allow, in particular, the determination of ages of maximal acceleration and maximal deceleration, in addition to the classical age of minimum and maximum velocity, both corresponding to zero acceleration. In

analogy to many processes in physics, acceleration might also be closer to the underlying dynamics than velocity (Gasser *et al.* 1983 b). An example is the higher intensity of the decelerating trend after the velocity peak during the PS, compared to the accelerating trend before it. This could also be inferred from the asymmetry of the velocity peak, but can be more readily seen and quantified (Gasser *et al.* 1983 a) in the acceleration curve.

There is no single method that satisfies all needs in curve fitting, and kernel estimation is not expected to be superior in every respect. The question, nevertheless, arises for what reason we did not use parametric fitting, which is the most common method. It has been postulated in the introduction that whenever the model (for distance) is not perfect, this will lead to more serious errors in terms of velocity and acceleration. In the analysis presented, this proved to be the case for a function as sophisticated as model 3 of Preece and Baines (1978). The lack of an MS in the model leads to serious discrepancies in the pre-adolescent period, both for boys and girls; they can be seen in the single examples and in the comparison with averaged raw velocity and acceleration (figures 7 and 8). Regarding the lack of fit below four years, it is fair to note that Preece and Baines (1978) evaluated their approach from four years onwards. The agreement of the parametric fit with the data, and with the kernel method, is better for the period of the PS, in particular regarding the peak. The onset of the PS (T6 etc.) is estimated at too early an age for parametric fitting, in particular for girls, confirming a finding by Hauspie, Wachholder, Baron, Cantraine, Susanne and Graffar (1980). As verified in Gasser *et al.* (1984), this bias for the PS is a consequence of the lack of an MS in the model, with more drastic consequences for girls for whom MS and PS follow in close succession (Gasser *et al.* 1983 a).

When discussing potential pitfalls of parametric fitting, it is worthwhile to reconsider the old problem — whether the logistic or the Gompertz function are better models for the PS (answered in favour of the logistic function by Marubini, Resele and Barghini 1971). The velocity peak of the PS is definitely asymmetric (see examples and averaged raw velocities and accelerations) and more so for girls (Gasser *et al.* 1983 a); this asymetry is reflected in a higher maximal deceleration than acceleration, with a difference that varies interindividually. The logistic function leads to a symmetric velocity peak (maximal acceleration = maximal deceleration); The Gompertz function has an asymmetric peak, but the wrong way round: maximal acceleration is larger than maximal deceleration. This provides one causal explanation for the inferiority of the Gompertz function. The logistic function, however, is also not fully adequate since it does not allow asymmetry, and does not account for the sex difference and interindividual variation in this respect.

Kernel estimation is intended to bring out the individual facets of growth without relying on a common model for all children, in which it has to be prespecified what pattern we expect (for example how many peaks in the velocity curve). The question is then how reliable this new method is and to what extent it can extract the finer features of the data without unduly following random variations. The latter point is particularly relevant for accelerations which are difficult to estimate, given measurements intervals as large as one year, and which are prone to follow random variation. The onset and the peaking of the PS which can be obtained from the individual velocity curve, show, however, an excellent agreement (in terms of mean, standard deviation and rank correlation) when extracted via acceleration. The comparison undertaken with averaged raw velocities and accelerations show that kernel estimates follow the data reliably across the sample of subjects. The evaluation of the choice of the smoothing parameter under-

taken in Gasser *et al.* (1984) has been deliberately on the conservative side; the individual velocity and, as a consequence, acceleration curves are expected to rather suppress random variation (possibly at the cost of obscuring true structure of the data). As a matter of fact, the velocity peak of the PS (VT8) for boys is on the average smaller by 0·39 cm/yr compared to parametric fitting, and this difference has to be interpreted as a bias effect. A closer statistical analysis (Gasser *et al.* 1984), however, revealed that both methods underestimate the size of the peak appreciably; it averages 9·77 cm/yr for boys (compard to 8·70 for parametric fitting and to 8·31 for the kernel method) and 7·89 cm/yr for girls (instead of 7·09 for parametric fitting and 7·00 for the kernel method).

The use of cubic smoothing splines is close to kernel estimation in spirit and results. In the methodological investigation (Gasser *et al.* 1984), cubic smoothing splines were somewhat inferior to kernels in terms of mean squared error and more expensive to compute; they are also not very attractive for estimating the acceleration curve which consists piecewise of straight lines between measurements. This led us to prefer kernel estimation and to abandon cubic smoothing splines.

After completion of our analysis, a variable knot cubic spline method was suggested by Berkey *et al.* (1983) for analysing height growth, and in particular for the MS. Variable knot cubic splines offer a more parsimonious approximation to the data compared to cubic smoothing splines. This method was judged by the authors to be inadequate for adolescent growth and also in the tails of the data. This might find an explanation in terms of the acceleration curve, indicating once more its utility; variable knot cubic splines result, as do cubic smoothing splines, in straight lines between knots: Berkey *et al.* (1983) allowed for five knots, with initial placements at 1·95, 6·5, 10·0, 13·5 and 20 years. (The numerical procedure of varying knots changes the position of knots but not their number and, therefore, not the number of piecewise linear functions in the acceleration curve.) This means, then, that the acceleration curve consists of one straight line between 10·0 and 13·5 and a further one between 13·5 and 20·0 years. An inspection of the examples and of the averaged raw accelerations presented shows that this provides an insufficient approximation to the data for the adolescent period and in post-adolescence. Berkey *et al.* (1983) found their method to model growth well between 5 and 10 years, as judged from a residual analysis in terms of distance; in terms of acceleration, however, the pre-adolescent fit consists of just two straight lines between 1·95 and 10·0 years when assuming the initial placement of knots, and this is probably not fully adequate for modelling the MS. We, therefore, feel that an analysis of the MS might be performed more advantageously by kernel estimation, using the acceleration curve in addition to the velocity curve.

Acknowledgements

This study was partially supported by the Fritz Hofmann-La Roche Foundation. The data analysis has been part of the research programme of the Sonderforschungsbereich 123 (project B1) at the University of Heidelberg and was made possible by financial support from the Deutsche Forschungsgemeinschaft.

References

BERKEY, C. S., REED, R. B., and VALADIAN, I., 1983, Midgrowth spurt in height of Boston children. *Annals of Human Biology*, **10**, 25–30.

COLLOMB, G., 1981, Estimation nonparamétrique de la regression: revue bibliographique. *International Statistical Revue*, **49**, 75–93.

DEUFLHARD, P., and APOSTOLESCU, V., 1980, A study of the Gauss-Newton method for the solution of nonlinear least squares problems. In *Special Topics of Applied Mathematics*, edited by J. Frehse, W. Pallaschke and U. Trottenberg (Amsterdam: North Holland), pp. 129–150.

EL LOZY, M., 1978, A critical review 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).

GASSER, Th., and MÜLLER, H. G., 1979, Kernel estimation of regression functions. In *Smoothing Techniques for Curve Estimation*, edited by Th. Gasser and M. Rosenblatt (Heidelberg: Springer-Verlag), pp. 23–68.

GASSER, Th., MÜLLER, H. G., KÖHLER, W., PRADER, A., LARGO, R., and MOLINARI, L., 1983 a, An analysis of the midspurt and of the adolescent growth spurt of height based on acceleration. (Mimeographed manuscript.)

GASSER, Th., KÖHLER, W., MÜLLER, H. G., LARGO, R., MOLINARI, L., and PRADER, A., 1983 b, Human height growth: correlational and multivariate structure, with implications for regulation. (Mimeographed manuscript.)

GASSER, Th., MÜLLER, H. G., KÖHLER, W., MOLINARI, L., and PRADER, A., 1984, Nonparametric regression analysis of growth curves. *Annals of Statistics*, **12**, 210–229.

GASSER, Th., and MÜLLER, H. G., 1984, Estimating regression functions and their derivatives by the kernel method. *Scandinavian Journal of Statistics*, (accepted for publication).

GASSER, Th., MÜLLER, H. G., and MAMMITZSCH, V., 1985, Kernels for nonparametric curve estimation. *Journal of the Royal Statistical Society*, (accepted for publication).

HAUSPIE, R. C., WACHHOLDER, A., BARON, G., CANTRAINE, F., SUSANNE, C., and GRAFFAR, M., 1980, A comparative study of the fit of four different functions to longitudinal data of growth in height of Belgian girls. *Annals of Human Biology*, **7**, 347–358.

LARGO, R. H., GASSER, Th., PRADER, A., STÜTZLE, W., and HUBER, P. J., 1978, Analysis of the adolescent growth spurt using smoothing spline functions. *Annals of Human Biology*, **5**, 421–434.

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**, 237–251.

MARUBINI, E., 1978, Mathematical handling of long-term longitudinal data. In *Human Growth*, edited by F. Falkner and J. M. Tanner, (New York: Plenum Press), vol. I, pp. 209–225.

MOLINARI, L., LARGO, R., and PRADER, A., 1980, Analysis of the growth spurt at age seven (mid-growth spurt). *Helvetica Paediatrica Acta*, **35**, 325–334.

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.

ROSENBLATT, M., 1956, Remarks on some nonparametric estimates of a density function. *Annals of Mathematical Statistics*, **27**, 832–835.

STÜTZLE, W., GASSER, Th., MOLINARI, L., PRADER, A., and HUBER, P. J., 1980, Shape-invariant modeling of human growth. *Annals of Human Biology*, **7**, 507–528.

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 Diseases in Childhood*, **41**, 454–471, 613–635.

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

Address correspondence to: Dr T. Gasser, Zentralinstitut für Seelische Gesundheit, P.O.B. 5970, D-6800, Mannheim 1, FR Germany.

Zusammenfassung. Es wird eine Methode zur Schätzung von Beschleunigung, Geschwindigkeit und Distanz in Längsschnitt-Wachstumskurven vorgestellt, und sie wird durch die Analyse des menschlichen Körperhöhenwachstums illustriert. Dieser Ansatz, genannt Kernel-Schätzung, gehört zur Klasse der Glättungsmethoden und nimmt nicht a priori ein festes Funktionsmodell an, und nicht einmal jenes und dasselbe Modell ist für alle Kinder anwendbar. Die vorgestellten Beispiele zeigen, daß Beschleunigungskurven eine bessere Quantifizierung des mittleren Wachstumsschubes (MS) und eine differenziertere Analyse des puberalen Schubes (PS) erreichen. Beschleunigung tendiert dazu, Zufalls-variationen der Daten zu folgen, und Parameter, die in Form von Beschleunigung definiert sind, werden daher durch einen Vergleich mit Parametern validiert, die in der Form von Geschwindigkeit definiert sind. Unser Ansatz einer nichtparametrischen Kurvenanpassung wird auch mit parametrischer Anpassung entsprechend einem Modell von Preece und Baines (1978) verglichen.

Résumé. Il est présenté une méthode d'estimation de l'accélération de la vitesse et de la distance de courbes de croissance longitudinales, qui est illustrée par une analyse de la croissance de la taille humaine. Cette approche, appelée estimation de noyau, appartient à la classe de méthodes de lissage et n'assume pas de modèle fonctionnel fixé *a priori*, même pas que un seul et le même modèle soit applicable à tous les enfants. Les exemples présentés montrent que les courbes d'accélération pourraient permettre une meilleure quantification de la poussée intermédiaire de croissance (MS) et une analyse plus différenciée de la poussée pubertaire (PS). Les accélérations sont sujettes à suivre des variations aléatoires présentes dans les données; les paramètres définis en termes d'accélération sont, par conséquent, validés par une comparaison avec des paramètres définis en termes de vitesse. Notre approche d'ajustement non paramétrique de courbe est aussie comparée avec un ajustement paramétrique selon un modèle suggéré par Preece et Baines.