



A New Generalized Logistic Sigmoid Growth Equation Compared with the Richards Growth Equation

COLIN P. D. BIRCH*

The Macaulay Land Use Research Institute, Craigiebuckler, Aberdeen AB15 8QH, UK

Received: 15 September 1998 Returned for revision: 22 January 1999 Accepted: 3 March 1999

A new sigmoid growth equation is presented for curve-fitting, analysis and simulation of growth curves. Like the logistic growth equation, it increases monotonically, with both upper and lower asymptotes. Like the Richards growth equation, it can have its maximum slope at any value between its minimum and maximum. The new sigmoid equation is unique because it always tends towards exponential growth at small sizes or low densities, unlike the Richards equation, which only has this characteristic in part of its range. The new sigmoid equation is therefore uniquely suitable for circumstances in which growth at small sizes or low densities is expected to be approximately exponential, and the maximum slope of the growth curve can be at any value. Eleven widely different sigmoid curves were constructed with an exponential form at low values, using an independent algorithm. Sets of 100 variations of sequences of 20 points along each curve were created by adding random errors. In general, the new sigmoid equation fitted the sequences of points as closely as the original curves that they were generated from. The new sigmoid equation always gave closer fits and more accurate estimates of the characteristics of the 11 original sigmoid curves than the Richards equation. The Richards equation could not estimate the maximum intrinsic rate of increase (relative growth rate) of several of the curves. Both equations tended to estimate that points of inflexion were closer to half the maximum size than was actually the case; the Richards equation underestimated asymmetry by more than the new sigmoid equation. When the two equations were compared by fitting to the example dataset that was used in the original presentation of the Richards growth equation, both equations gave good fits. The Richards equation is sometimes suitable for growth processes that may or may not be close to exponential during initial growth. The new sigmoid is more suitable when initial growth is believed to be generally close to exponential, when estimates of maximum relative growth rate are required, or for generic growth simulations.

© 1999 Annals of Botany Company

Key words: Asymptote, *Cucumis melo*, curve-fitting, exponential growth, intrinsic rate of increase, logistic equation, maximum growth rate, model, non-linear least-squares regression, numerical algorithm, point of inflexion, relative growth rate, Richards growth equation, sigmoid growth curve.

INTRODUCTION

Classical growth models, such as the logistic, Gompertz and Richards equations continue to be widely and frequently used to describe various biological processes (Werker and Jaggard, 1997). Many of these equations define sigmoid curves, in which the rate of growth increases as size increases from low values, reaches a maximum at a point of inflexion and then decreases towards zero at an upper asymptote, so that they look like the central part of a rotated S (Ratkowsky, 1983). Much current statistical software specifically assists the convenient fitting of such equations to observations. Most often biologists have used models like the logistic, Gompertz or monomolecular functions, which have only three parameters (if it is assumed that the minimum value is 0) (Zeide, 1993). These three parameters are usually an upper asymptote, a rate parameter and a time constant. The rate parameter determines the rate at which growth initially accelerates. The time constant determines the time at which the function has a specific value between its minimum and maximum, or the time when growth rate is maximum (e.g. Pienaar and Turnbull, 1973).

Some three-parameter growth functions, such as the Weibull function, do not need a time constant because they have zero value at time zero, so that they are more flexible in other respects (Yang, Kozak and Smith, 1978). However, because the curves defined by these equations are severely constrained, many have chosen to use equations with additional parameters. The Richards equation continues to be the most popular of these more flexible growth equations since it was first proposed for this purpose by Richards (1959). It is still used for diverse purposes, including modelling tree growth, the growth of juvenile mammals and birds, and comparisons of treatment effects on plant growth (Pienaar and Turnbull, 1973; Huang and Titus, 1994; Verwijst and Vonfircks, 1994; Parresol, 1995; Rennolls, 1995; Gaillard *et al.*, 1997).

The Richards equation has been popular for several reasons. It has an additional parameter, which is a shape parameter that can make the Richards equation equivalent to the logistic, Gompertz, or monomolecular equations (France and Thornley, 1984). Varying the shape parameter allows the point of inflexion of the curve to be at any value between the minimum and the upper asymptote. An important factor in its use instead of other sigmoid equations may be its availability as one of the standard equations

* Fax +44 (0) 1224 311556, e-mail C.Birch@mluri.sari.ac.uk

TABLE 1. Symbols used in equations in the text

Symbol	Meaning
y	A variable representing the value of a measure of size or density of an organism or population.
t	A variable representing time.
$f()$	An unspecified function.
a	The maximum intrinsic rate of increase (RGR) of y . Dimension equal to time ⁻¹ .
K	The upper asymptote of y . Dimension equal to size.
b	An additional parameter in the Richards equation introduced as a power law so that it can define asymmetric curves. Dimensionless scalar.
d	A parameter in the Richards equation which allows the time at which $y = K/2$ to be varied. Dimensionless scalar.
t_0	The time at which $y = K/2$.
c	An additional parameter in the new sigmoid equation introduced so that it can define asymmetric curves. Dimensionless scalar.
e	A universal constant, the base of the natural logarithm.
k	Lower asymptote of a logistic curve (Appendix).
z	The rate parameter of the upper of two logistic curves in one of the hybrid curves used to generate the simulated dataset. Dimension equal to time ⁻¹ (Appendix).

offered for curve-fitting by statistical software packages (e.g. Genstat 5 Committee, 1993; Sigmaplot® 4.0, 1997).

However, there have been criticisms of the Richards equation. Some authors consider that the additional parameter used in it has no obvious biological interpretation (Thornley and Johnson, 1990; Zeide, 1993). It has been accused of being so unstable numerically that its parameter estimates become 'useless' (Zeide, 1993), or 'practical problems are often encountered when fitting to experimental data' (Thornley and Johnson, 1990). On the other hand, well-written specialist applications of the Richards equation can be very stable (e.g. Genstat 5 Committee, 1993). Often it is reported as failing to provide good fits to observations, or being worse than alternative models (e.g. Brown and Mayer, 1988; Meng *et al.*, 1997; Werker and Jaggard, 1997). There are many alternative sigmoid curves, and some reviewers have advised against use of the Richards equation (Ratkowsky, 1983; Thornley and Johnson, 1990; Zeide, 1993; Shvets and Zeide, 1996).

In this paper, a new sigmoid growth equation with unique properties will be presented that has many of the features that make the Richards equation popular for curve-fitting. It can have a point of inflexion at any value between its minimum and its upper asymptote. In addition, it is more consistent than the Richards equation, because its minimum is always an asymptote, which makes the new equation more suitable for generic simulation models. The new sigmoid equation can be justified using the same theoretical arguments as the Richards equation, and it is similarly a generalization of the logistic equation.

Curve fits using the new sigmoid equation were compared with fits using the Richards growth equation. To avoid confounding this comparison with the biological peculiarities and error structures of particular real life examples, a dataset was simulated with known charac-

teristics and error structure. The two growth models could then be compared not only in terms of their ability to statistically fit the data, but also in terms of their ability to describe the known characteristics of the original data. The two models were also compared using the observations of Pearl, Edwards and Miner (1934), which were originally used by Richards (1959) in his presentation of his growth equation, so as to confirm that conclusions from the comparison made using simulated data are relevant to real life applications.

AN EMPIRICAL MODEL OF VEGETATION GROWTH

This study results from a requirement for an empirical equation to represent the growth of various individual species within mixed vegetation. It is likely that, at sufficiently low densities (i.e. mass per unit area), the growth (i.e. increase in mass) of an individual species will tend to be proportional to its mass. Pasteur's Principle of Homogenesis demands that there should be no potential for growth at zero mass (Huxley, 1870). Therefore all derivatives should be zero for zero mass. This leads directly to an exponential model of growth:

$$\frac{dy}{dt} = f(y)y \quad (1)$$

where y is a measure of the density (mass per unit area) of a species, t is time and $f(y)$ equals the relative growth rate (RGR), which reduces as y increases. Here we are not presenting a model of inter-specific competition, so we assume a monoculture for the rest of this presentation.

More or less the simplest non-constant form of $f(y)$ generates the logistic equation:

$$\frac{dy}{dt} = ay \left(1 - \frac{y}{K}\right) \quad (2)$$

where a is a constant equal to the RGR at very low values of y , and K is the maximum observable value of y . Although the logistic growth equation generates a curve that tends towards an exponential form at low values, its maximum slope, or 'point of inflexion', is always at half the value of the upper asymptote, $K/2$. This is unsatisfactory, because the factors that determine the density at which each species grows fastest are complex, so it is unlikely that all species grow fastest in monoculture when they are at half their maximum standing crop.

Introducing another parameter into the equation could allow the shape of the upper part of the curve to be independent of the shape of the lower part, while still having an equation that tends towards an exponential form at low values of y . One option is the Richards growth equation:

$$\frac{\partial y}{\partial t} = ay \left[1 - \left(\frac{y}{K}\right)^b\right] \quad (3)$$

where b is a constant that allows the shape of the sigmoid curve to be varied. When $b = 1$ the Richards equation

matches the logistic equation, but for $b > 1$ the maximum slope of the curve is when $y > K/2$, and when $b < 1$ the maximum slope of the curve is when $y < K/2$. This allows a wider range of curves to be produced, but as b tends towards zero, the lowest value of y at the point of inflexion remains greater than K/e , where e represents the universal constant, the base of the natural logarithm. In fact, as b tends towards zero the Richards growth curve tends towards the Gompertz growth curve, which has its steepest slope at $y = K/e$, and does not tend towards an exponential form for low y (Richards, 1959). For these reasons the integrated form of the Richards equation has two forms, depending on the value of b (Causton and Venus, 1981):

$$y = K(1 + e^{(d-abt)})^{-1/b} \quad \text{for } b > 0 \text{ and } a > 0 \quad (4)$$

$$y = K(1 - e^{(d-abt)})^{-1/b} \quad \text{for } 0 > b \geq -1 \text{ and } a < 0 \quad (5)$$

where d is a parameter that indirectly defines the value of t (time) at which $y = K/2$. It should be noted that eqn (4) tends towards an exponential equation at low values of y , and is the same as the logistic equation when $b = 1$. Hence the Richards equation has sometimes been called a generalized logistic equation. However, eqn (5) is intermediate between the monomolecular equation ($b = -1$) and the Gompertz equation (as b tends to 0). When $abt = d$, $y = 0$, but $dy/dt > 0$, so this equation predicts non-zero growth at zero mass, infringing the Principle of Homogenesis. The maximum RGR at low values of y is a in eqn (4), but a is negative in eqn (5) and the maximum RGR at low values of y is therefore undefined.

In addition to the maximum RGR and upper asymptote, various other characteristics of a growth curve described by the Richards equation can be derived from its parameters. For example:

$$\text{the maximum growth rate} = Kab(1+b)^{-(1+b)/b} \quad (6)$$

$$\text{the value of } y \text{ at which growth is maximum} = K(1+b)^{-1/b} \quad (7)$$

$$\begin{aligned} \text{the time at which } y = K/2 &= \frac{d - \ln(2^b - 1)}{ab} \quad \text{for } b > 0, \\ &= \frac{d - \ln(1 - 2^b)}{ab} \quad \text{for } b < 0 \end{aligned} \quad (8)$$

The Richards equation is unsuitable for constructing simple empirical simulations of the growth of vegetation because only eqn (4) satisfies our assumption that growth is near exponential at very low densities. For example, swards of some cultivars of ryegrass (*Lolium perenne* L.) can achieve their maximum net production at a mass density of only 200 g m⁻² (Alberda and Sibma, 1968). This is about 0.2 of the maximum standing crop of about 1000 g m⁻². Equation (4) only applies if the point of inflexion (equivalent to maximum net production) is above 0.368 (1/e) of the value of the upper asymptote. Equation (5) would have to be used for *Lolium perenne*, but it has intolerable properties, including non-zero growth at zero mass. For a generic simulation model, we require a sigmoid equation that can have its point of inflexion at any value, while always tending

towards exponential growth at low values. This is not just to satisfy biological preconceptions, but also because non-zero growth at zero mass is likely to cause a simulation to fail ('crash') during runs.

There are many other sigmoid equations, but very few can be used as generalizations of the logistic equation because they do not simulate exponential growth at small sizes or low densities. For example, the Weibull equation is excellent for empirical curve-fitting, and it ensures that size or density is zero at time zero, which is useful in many applications (Yang *et al.*, 1978; Ratkowsky, 1983; Brown and Mayer, 1988). However, although its slope at zero size or density may be zero, it predicts growth from zero mass and is therefore fundamentally different from the logistic equation.

A new generalized logistic equation

If the integrated form of a logistic equation is rearranged to solve for t , we obtain:

$$t - t_0 = \frac{\ln(y)}{a} - \frac{\ln(K-y)}{a} \quad (9)$$

where t_0 is the time at which $y = K/2$. One way to increase the flexibility of this equation would be to introduce a new parameter to give:

$$t - t_0 = \frac{\ln(y)}{a} - \frac{c \ln(K-y)}{a} + \frac{(c-1) \ln(K/2)}{a} \quad (10)$$

where c is a new shape parameter adjusting the relative importance of the acceleration of growth as y increases from low values and the deceleration of growth as y approaches K . The closeness of the relationship between this new equation and the original logistic equation means that it represents a 'generalized logistic equation' just as much as the Richards equation. The differential form of this equation, predicting growth rate as a function of size is:

$$\frac{dy}{dt} = \frac{ay(K-y)}{K-y+c} \quad (11)$$

This new sigmoid equation has several immediate advantages. Varying the value of c from 0 upwards allows the maximum rate of growth to be at any value of y from 0 to K . The equation is equivalent to the logistic equation when $c = 1$ and the simple exponential equation when $c = 0$. When $c < 1$, the maximum rate of growth is at $y > K/2$; when $c > 1$, the maximum rate of growth is at $y < K/2$. The two assumptions, that growth tends towards exponential growth at low y and that K is an upper asymptote for y , are obeyed for any positive value of c . The full range of sigmoid curves can be generated using only positive values of the parameters a , c and K . Furthermore, the absence of a power law is an important advantage for numerical application. The only disadvantage of this equation is that it cannot be integrated to give an analytical solution for y . However, almost all applications of growth equations are numeric, so this is a minor problem.

Like the Richards equation, several other characteristics of a growth curve described by the new equation can be derived from its parameters:

the maximum RGR = a for all positive values of c

$$\text{the maximum growth rate} = \frac{aK}{(\sqrt{c}+1)^2} \quad (12)$$

$$\text{the value of } y \text{ at which growth is maximum} = \frac{K}{\sqrt{c}+1} \quad (13)$$

MATERIALS AND METHODS

Simulated dataset

The new sigmoid equation was expected to have advantages over the Richards equation when representing growth processes that tend towards exponential growth at small sizes or low densities. These advantages were expected to be most significant when the point of inflexion was substantially below half the value of the upper asymptote. Construction of the test dataset therefore required a sigmoid curve that: (1) tended towards exponential growth at low values; (2) had points of inflexion at diverse positions relative to the upper asymptote; (3) did not have any other intrinsic reasons for being more easily fitted by one equation than the other.

The logistic equation tends towards exponential growth at low values of its dependent variable and is a special case of both the Richards equation and the new sigmoid equation, satisfying conditions (1) and (3). Condition (2) was satisfied by splining together different logistic curves, so that the resulting sigmoid hybrid curves had points of inflexion at whatever values were required (see Appendix).

To simplify analysis, the hybrid curves were standardized as much as possible. The lower part of every hybrid curve was the lower part of a logistic with a maximum RGR [a in eqn (2)] equal to 1 and a lower asymptote equal to zero; the upper part was equal to the upper part of a logistic curve with the equivalent parameter a equalling a value from the series 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 0.75, 1, 1.5, 2, 5. There were therefore 11 different hybrid curves. One of the hybrid curves, in which the parameter a of both its upper and lower parts equalled 1, was the logistic growth curve with maximum RGR equal to 1. The pairs of logistic curves were joined so that the hybrid curves had the following properties: (1) the lower asymptote for every hybrid curve was zero; (2) the upper asymptote for every hybrid curve was 100; (3) the two parts of each hybrid curve joined at its steepest point (point of inflexion). Note that this was usually the steepest point along only one of the two logistic curves used to generate it; (4) the joins were at time zero; (5) the two parts of each hybrid curve had equal value and slope where they joined, so that the value and slope of the hybrid curves were continuous.

The equations used for the eleven curves have been provided in the Appendix, which also demonstrates that the hybrid curves were continuous with continuously changing slope. The curves and the test dataset generated from them

were intended to be abstract, but for ease of reference the vertical axis is referred to as 'size', while the horizontal axis is time (Fig. 1).

Robust estimation of all the parameters in the two sigmoid equations required several points near the upper asymptote, several at small sizes and an adequate sample near the point of inflexion. Therefore twenty original points were chosen along each hybrid curve, ten before the point of inflexion at time zero and ten after (Fig. 1). The first point of every series was where size = 0.1. The next nine points were chosen to be at regular time intervals, so that the interval between the tenth point and time zero was half the interval between adjacent previous points. Along curves in which the point of inflexion was equal to or below size = 50, the last point was where size = 99.9 (100 - 0.1). Points 11 to 19 were chosen at regular time intervals between time = zero and the last point, so that the interval between time = zero and the eleventh point was half the interval between adjacent following points. Thus the 20 original points along the logistic curve were at regular time intervals between the first at size = 0.1 and the last at size = 99.9. For the remaining curves, this approach would have selected points too close in time to allow accurate estimation of the upper asymptote. Therefore, along curves in which the point of inflexion was above size = 50, points 11 to 20 were chosen to be at the same regular time intervals as used for the first ten points.

The basis of the simulated dataset was therefore 11 sets of 20 points each. The comparison was supposed to test the suitability of the new sigmoid equation and the Richards equation for fitting growth processes that tended towards exponential growth at small sizes. It was therefore appropriate to logarithmically transform all points, so as to emphasize growth at small sizes.

From each of the 11 sets of 20 logarithmically-transformed points, 100 perturbed replicates were generated by adding randomly generated error terms to each point. Each error term was generated from a normal distribution with mean = zero and standard deviation = 0.09531 [$\log_e(1.1)$], using a random number generator (Maple V Release 4; Redfern, 1993). The error terms were therefore approximately equivalent to a coefficient of variation = 10%. Adding normally distributed error terms with uniform variance to a logarithmically transformed dataset meant that the errors in the untransformed sizes were proportional to the original sizes at each point. Current frequent use of logarithmic transformations in statistical analysis of biological measurements suggests that biological variables that cannot have values below zero often have standard deviations proportional to their means.

The construction of the curves defined the values of various characteristics, which could be compared with the estimates from fitting the two sigmoid equations. The maximum RGR of all curves was 1.0, the upper asymptote was 100. The time t_0 at which size = $K/2$, the maximum slope (growth rate) and the size at the point of inflexion could all be calculated from the logistic equations that were used to generate the hybrid curves (Appendix). These values were compared with those calculated from the parameters of fitted Richards equations, using eqns (3) to (8), and from

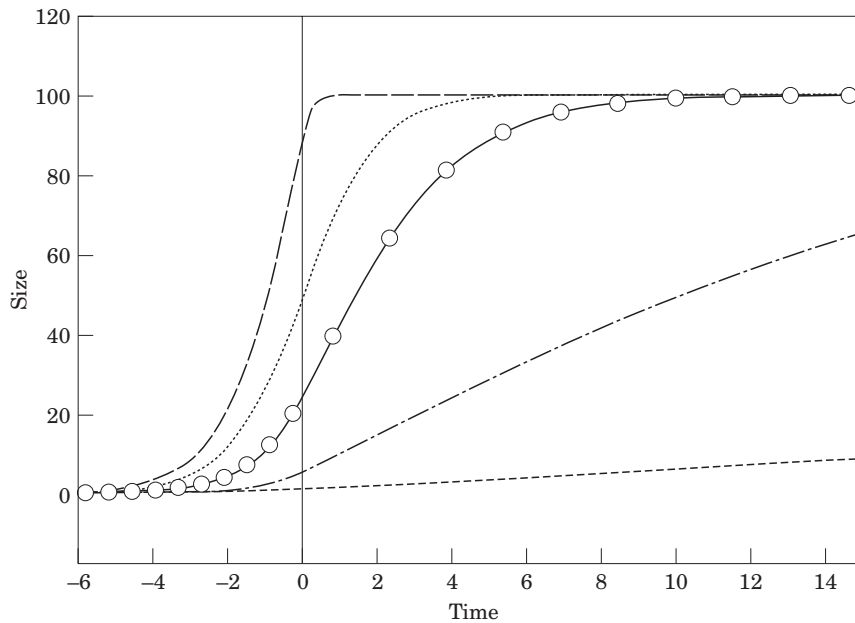


FIG. 1. Five examples from the 11 hybrid curves used to generate the test data for comparing the new sigmoid equation with the Richards equation. The curves are labelled in the key according to the value of the rate parameter [a in eqn (2)] of the logistic equation used for the upper half of each curve. The range of values for this parameter among all 11 curves was 0.01–5.0, so this graph includes the two most extreme curves. The curve with a value of 1.0 is the curve of the logistic equation. All curves have their points of inflexion at time = 0. ○, 20 points along the 0.5 curve that were used to generate the test dataset for that curve by adding random, lognormally-distributed errors. Note that ten of the 20 points are below the point of inflexion and ten are above. Each set of ten points is regularly spaced along the time axis, as explained in the text. Rate parameter used for upper half of each curve: -- 5.0; 1.0; — 0.5; -.- 0.1; --- 0.01.

the parameters of fits of the new sigmoid equation, using eqns (10) to (13).

Observed dataset

Richards (1959) demonstrated his proposed growth equation by applying it to observations of Pearl *et al.* (1934) on the height growth of *Cucumis melo* seedlings at six different temperatures. Here the Richards equation and the new sigmoid equation are compared using the dataset originally selected by Richards. Since these were means of non-destructive measurements, fits were made to the untransformed observations. Preliminary analysis indicated that deviations from the curve fits did not increase with the estimated values.

Curve-fitting

Symbolic computation software (Maple V Release 4; Redfern, 1993) was used to evaluate the Richards equation exactly from eqns (4) and (5) and to evaluate the new sigmoid by solving eqn (10) for y numerically. The solution to eqn (10) is bracketed by:

$$\text{for } c > 1.0 \quad \frac{Kx}{1+cx} < y < \frac{Kx}{1+x}$$

$$\text{for } c < 1.0 \quad \frac{Kx}{1+x} < y < \min\left(\frac{Kx}{1+cx}, K\right) \quad (14)$$

where $x = 2^{c-1}e^{\alpha(t-t_0)}$. An exact solution for y can be obtained immediately for c = 1. A precise solution for y was then obtained using a combination of a bisection method and a Newton-Raphson method (Press *et al.*, 1992). Numeric integration might have been a faster alternative approach. The least-squares best fits to each sequence of points for the parameters of the two equations were found by a purpose-built minimization programme using symbolic computation software (Redfern, 1993). The algorithm was derived from the Nelder and Mead downhill simplex method described in Press *et al.* (1992). Statistical analysis of the results of curve-fitting was carried out in Maple V Release 4 and on a computer spreadsheet (Redfern, 1993; Microsoft®, 1996). All curve fits were repeated at least once to achieve confidence that a true best fit had been achieved by the least-squares criterion.

RESULTS

Simulated dataset

After log-transformation, the average sum of squared deviations of each sequence of 20 points from the original curve that they were derived from was 0.1817. The average sum of squared deviations for fits of the new sigmoid equation to each of the eleven curves was less than this (Fig. 2A). The average sum of squared deviations for fits by the new sigmoid equation was also lower than for the Richards equation for every curve, although the difference was very small when the rate parameter used to generate the upper part of the curve was greater than or equal to 0.5 (Fig. 2A). When the rate parameter of the upper part of the curve was

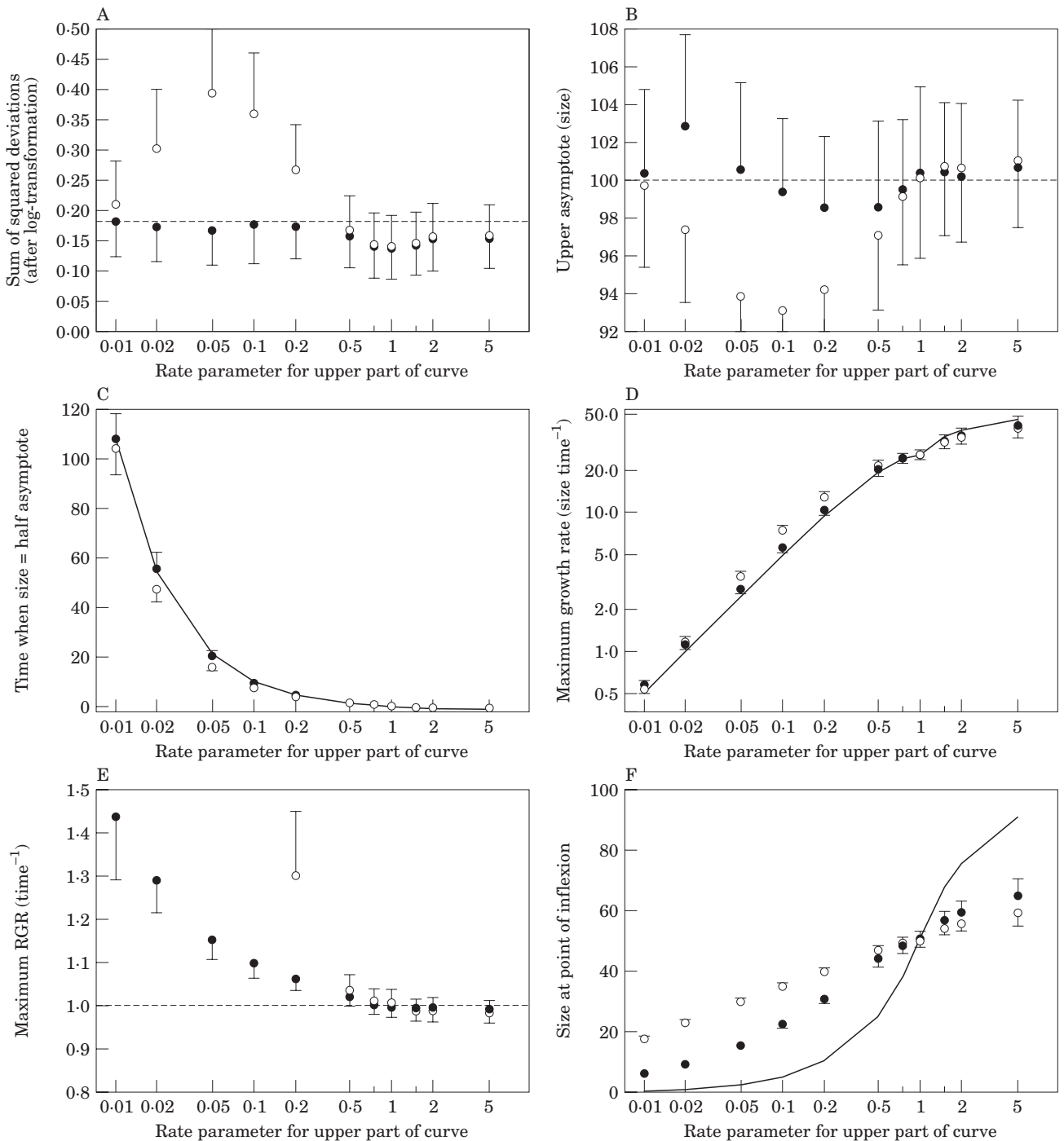


FIG. 2. Results of fitting the Richards equation (○) and the new sigmoid equation (●) to sets of curves randomly generated from the 11 curves partly illustrated in Fig. 1. Points are means of 100 estimates. Bars = s.d. A, Sum of squared deviations of fitted curves: the horizontal dashed line indicates expected average deviation from the original curve used to generate each set. B, Estimated values of the upper asymptote: the horizontal dashed line indicates original value of 100. C, Estimated time when size reaches half of its maximum value: the solid line indicates original time for each curve. D, Estimated maximum rate of growth: the solid line indicates original maximum rate of growth of each curve. Vertical axis is log-transformed. E, Estimated maximum RGR: the horizontal dashed line indicates the original value of 1.0 for every curve. Four points have been omitted from the Richards equation's estimates because they were either undefined or greater than 50. F, Estimated size at which growth rate is maximum: the solid line indicates original size for each curve.

less than or equal to 0.2, the Richards equation consistently fitted badly. The worst fits were when the rate parameter of the upper part of the curve was 0.05.

This pattern was repeated for estimates of the various characteristics of the curves (Fig. 2B–F). The estimates from the new sigmoid equation always tended to be closer

TABLE 2. Sum of squared deviations from fitted curves and estimates of characteristics of height growth by *Cucumis melo* seedlings in darkness at six different temperatures (Pearl et al., 1934)

Treatment Temp. (°C)	Sum of squared deviations		Time for 1/2 maximum height (d)		Maximum height (mm)		Proportion of maximum height at point of inflexion		Maximum elongation rate (mm d ⁻¹)		Maximum RGR (d ⁻¹)	
	R	New	R	New	R	New	R	New	R	New	R	New
37	3.95	8.06	4.60	4.60	157	157	0.398	0.391	36.2	37.1	3.93	1.54
35	17.2	26.5	4.22	4.22	196	197	0.377	0.372	48.2	49.5	14.2	1.82
30	106	139	4.98	4.99	249	249	0.368	0.373	54.9	56.3	X	1.63
25	38.6	51.5	5.27	5.27	226	227	0.383	0.374	50.6	52.1	7.52	1.64
20	15.7	14.7	8.93	8.93	178	179	0.455	0.447	32.9	33.3	1.10	0.93
15	1.52	0.89	25.3	25.3	74.1	73.9	0.551	0.563	7.5	7.4	0.29	0.32
	N.S.		N.S.		N.S.		*		†		‡	

Columns marked ‘R’ are values estimated using the Richards equation; columns marked ‘New’ are values estimated using the new sigmoid equation. The symbol ‘X’ marks an undefined value of RGR. NS, No significant difference found between the two methods ($P > 0.05$); *, the Richards equation estimates points of inflexion closer to 0.500 ($P < 0.05$, one-tailed t -test); †, elongation rates estimated by the new sigmoid equation are higher (mean difference = 0.9 mm d⁻¹, $P < 0.05$, two-tailed t -test); ‡, variance of RGR among treatments from Richards equation = 8.90; variance from new sigmoid equation = 0.33.

to the correct values than estimates from the Richards equation. Nevertheless, although worse than the new sigmoid equation, the Richards growth equation gave good estimates for most of the characteristics of curves in which the rate parameter used to generate the upper part of the curve was greater than or equal to 0.5 (Fig. 2B–E). For the other curves, the Richards equation made inaccurate but possibly acceptable estimates of the maximum size, the time at which size reached half its maximum ($t_{0.5}$) and the maximum growth rate (Fig. 2B–D). When the upper rate parameter was in the range 0.02 to 0.2, the error in the Richards equation’s average estimate of each characteristic was often larger than the standard deviation among estimates. In general the coefficients of variation of the sets of estimates of each characteristic tended to be less than 10%, so the average error or bias was a substantial part of the error in each estimate (Fig. 2B–F).

The Richards equation’s worst performance was in estimating the maximum RGR (Fig. 2E). Its average estimate was 1.3 (correct value 1.0) for the curve generated with an upper rate parameter of 0.2. When the upper rate parameter was reduced even further to 0.1, estimates of the maximum RGR became almost random, varying between 1.5 and 383 (not included in Fig. 2E). For the remaining curves, which were even more asymmetric, the maximum RGRs were undefined. The new sigmoid estimated the maximum RGRs of these curves less well than other characteristics, although its estimates were mainly within 50% of the correct value for the most asymmetric curve and better for the other curves.

Estimates of the size at the point of inflexion tended to be closer to 50 than the true value (Fig. 2F). The Richards equation always gave the estimates most distant from the true value, so that its average estimates for the 11 curves ranged only from 17.8 to 59.0, compared with the true values that ranged from 0.5 to 90.0. The Richards equation gave poor estimates of this characteristic for all curves except those very close to a simple logistic, even though the coefficient of variation of each set of estimates was usually

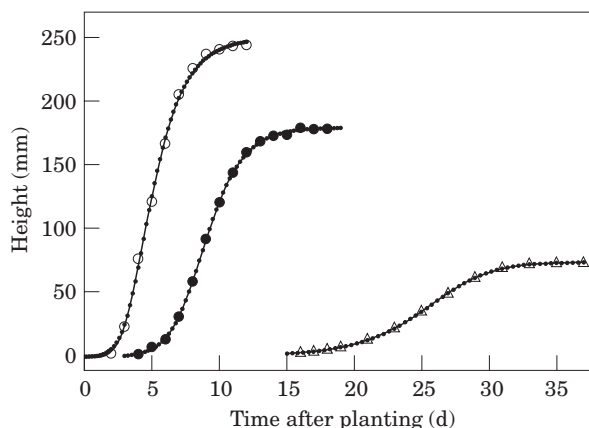


FIG. 3. Height growth in the dark of seedlings of *Cucumis melo* at three different temperatures: 30 °C (○), 20 °C (●) and 15 °C (△) (data from Pearl et al., 1934). Curves were fitted using the Richards equation (solid line) and the new sigmoid equation (dotted line). The curve fits are only visibly different near the base of the 30 °C curve.

less than 5%. The estimates from the new sigmoid ranged from 6.4 to 64.4.

Observed dataset

The Richards equation fitted the height growth at the four warmest temperatures more closely than the new sigmoid equation, but both equations gave excellent fits and there was no statistically significant evidence of a consistent difference between methods in the sum of squared deviations (Table 2, Fig. 3). The two equations estimated similar values for most characteristics of the curves (Table 2). For the four warmest temperatures, both equations estimated the height when elongation rate was maximum to be close to 0.37-times the maximum height. Nevertheless, there was evidence that the Richards equation tended to estimate that the point of inflexion was closer to half the maximum height than did the new sigmoid (Table 2). This result was associated with

a small difference in the estimated maximum elongation rate. The largest difference between the two methods was in the variance among temperature treatments of the maximum RGR of seedling height (Table 2), and the Richards equation did not estimate the maximum RGR at 30 °C.

DISCUSSION

The greater theoretical consistency and the more generic nature of the new sigmoid model may in themselves be sufficient reason for choosing to use it in preference to the Richards equation for various simulations and curve-fitting problems. However, since it is essentially an empirical model, it was appropriate to test whether these theoretical advantages would result in noticeable benefits in practice. These tests also indicated the types of growth processes that may be most appropriately represented by the Richards equation or the new sigmoid equation.

Comparison of curve fits

The new sigmoid equation fitted all 11 sets of curves in the simulated dataset successfully (Fig. 2). On average, for every set, it found a curve that fitted each series of 20 points more closely than the curve that originally generated them. Given the wide range of forms of the original curves, this demonstrated the versatility of the new equation (Fig. 1). Its average estimates of maximum size and the time at which size reached half its maximum were always within about 3% of the correct values (Fig. 2B and C). The coefficients of variation of estimates were usually less than 10%. Its average estimates of maximum growth rate were slightly biased, but were always within 15% of the correct values (Fig. 2D). Only its estimates of the size at the point of inflexion were inaccurate, although closer than estimates from the Richards equation (Fig. 2F).

The new sigmoid also made good estimates of the maximum RGR for curves that were only slightly asymmetric, or which had the point of inflexion above half the maximum size (Fig. 2E). Even for curves with points of inflexion at very small sizes, the estimates of maximum RGR remained reasonably close to the correct values. However, these encouraging results depended on the very small error terms at small sizes, which resulted from errors being proportional to size. Often such a relationship between the magnitude of measurement errors and size will not be maintained down to sizes as small as 0.1% of the upper asymptote.

In contrast, the Richards growth equation gave consistently worse fits to the simulated dataset and less accurate estimates of the characteristics of the original curves (Fig. 2). It made especially inaccurate estimates of the characteristics of curves with points of inflexion substantially below half the maximum size. The estimates of maximum RGR were particularly unreliable, except when the point of inflexion was near or above half the asymptotic value.

Both equations made estimates of the size at the point of inflexion that were heavily biased towards half the maximum value of size, although this trend was weaker for the new sigmoid than the Richards equation (Fig. 2F). This is a

problem that deserves attention, because the results of curve-fitting are sometimes summarized using this parameter (e.g. Verwijst and Vonfircks, 1994; Gaillard *et al.*, 1997). These results suggest that estimates of this parameter derived by curve-fitting may be inaccurate and may underestimate the degree of asymmetry. However, the discontinuity of the second derivative of each hybrid curve near the point of inflexion may have influenced these results. Curve fits to the experimental observations also indicated that the Richards equation tended to estimate that curves were more symmetric than the new sigmoid equation did.

Both equations obtained excellent fits to the experimental observations (Pearl *et al.*, 1934). The Richards equation's failure to consistently estimate maximum RGR at high temperatures can be justified by arguing that initial height growth was probably not exponential at these temperatures. In fact growth of these seedlings was close to what would be predicted by the Gompertz equation, with the point of inflexion close to $1/e$ (approx. 0.368) times the maximum height. This explains the very low residual sums of squares from the fits of the Richards equation, which includes the Gompertz equation as a special case. On the other hand, the new sigmoid's consistent plausible estimates of maximum RGR suggest that it can produce robust estimates of this parameter, even when particular sequences of observations appear to deviate from its assumption that growth tends towards exponential at small sizes. Moreover, the new sigmoid provided estimates of the size at the point of inflexion at high temperatures that were very close to the position expected for the Gompertz curve, so that careful interpretation of its curve fits could have detected the close match to the Gompertz curve.

The theoretical argument

The theoretical justification for both equations is strongest when they are used to represent growth processes that are expected to be exponential at small sizes or low densities. Studies of such processes often treat the intrinsic rate of increase (or RGR) of the initial exponential growth as a characteristic of particular populations, conditions or genotypes (e.g. Grime, Hodgson and Hunt, 1988; Thornley, 1995; Birch and Shaw, 1997). Indeed maximum RGR is used in plant growth analysis to characterize seedling growth, even though initial growth is only approximately exponential (Grime *et al.*, 1988; Poorter and Remkes, 1990). In such cases, the suitability of a growth equation for estimating maximum RGR is an important part of its function. Therefore the numerical problems in estimating parameters for the Richards equation also undermine its theoretical justification.

The theoretical argument supports the new sigmoid equation more consistently than the Richards equation. For all values of the additional parameter, the new sigmoid equation tends towards an exponential equation at small sizes or low densities. All the curves it produces can therefore be justified by the same theoretical argument. The Richards equation can only be supported by the theoretical argument when the parameter $b > 0$ [eqn (4)]. Since it is not always obvious before curve-fitting whether or not $b > 0$,

this undermines the theoretical justification for the Richards equation generally as well as for $b < 0$.

Use of the Richards equation

The Richards equation is generally flawed by its inconsistent properties. In part of its range it is exponential at small sizes or low densities; in the rest of the range it is not exponential in any respect. This means that parameters of different curves fitted using the Richards growth equation are not necessarily equivalent. In particular, one parameter equals the maximum intrinsic rate of increase (RGR) for some curves, but not others. This inconsistency undermines the comparability that is usually one of the main reasons for performing curve-fitting. In a simulation, it could also lead to unexpected behaviour that is difficult to interpret. Furthermore, Ratkowsky (1983) concluded that curve-fitting with the Richards equation was very likely to generate biased parameter estimates.

On the other hand, the Pearl *et al.* (1934) dataset illustrates the possibility that some phenomena may tend towards exponential growth at small sizes in some conditions, but have non-exponential growth at small sizes in other conditions. This could be because different processes limit growth in different conditions. When growth at small sizes may be limited either by a process that can be described by the Gompertz or monomolecular equations, or by a competing process that resembles exponential growth, the Richards equation may be very suitable.

Both the Richards equation and the new sigmoid equation could include additional parameters allowing the lower asymptote to have a non-zero value (e.g. Genstat 5 Committee, 1993; Sigmaplot® 4.0, 1997). This option was ignored here, because a lower asymptote not equal to zero would be inconsistent with the assumption that initial growth was exponential. The error terms used in the construction of the simulated dataset would also have been difficult to justify if lower asymptotes had been allowed to be non-zero. However, animal egg and birth masses, and, less frequently, plant seed masses are often a significant proportion of the organisms' maximum masses. Note that these initial masses and the dynamics of growth shortly after hatching, birth, or germination may substantially undermine assumptions that initial growth is close to exponential. Since the Richards equation does not necessarily represent exponential growth at small sizes, it may be more suitable than the new sigmoid equation when a non-zero lower asymptote is expected.

Use of the new sigmoid equation

Apart from not having a simple analytic integrated form of the equation available, the new sigmoid equation seems to have many advantages. Although many sigmoid equations have been described in the past (e.g. Zeide, 1993), this one is unique because it is the first that combines exponential growth at small sizes with the potential to have its point of inflexion at any point. It can produce a complete

range of sigmoid curves with both upper and lower asymptotes. The parameters generated by fitting curves to different sequences of observations can be directly compared and can be related to biologically meaningful characteristics of growth processes. In addition to applications in curve-fitting, it is also a useful basis for a simple, flexible and generic model of growth for inclusion in simulations of processes such as the vegetation dynamics of grazed plant communities (Birch, 1995; Birch, Werkman and Partridge, 1997).

In this paper, the two sigmoids were compared using good datasets that included points near both lower and upper asymptotes, as well as points close to the point of inflexion. Actual observations may often be less complete. In such cases, it will often be difficult to obtain separate estimates of all the parameters in these sigmoid equations. More constrained models must be used for fitting, such as logistic or linear models. For example, if observations of the increase in mass y are only available for small masses, the maximum RGR a can be estimated by estimating the slope of a regression of $\ln(y)$ against time. The parameters in fits of the new sigmoid equation were highly consistent with the parameters in the logistic equations used to generate the simulated dataset. This suggests that parameters for simulations using the new sigmoid equation can be obtained from multiple sources, when information is only available as small fragments. The direct relationships between parameters of the new sigmoid equation and measurable characteristics of biological growth suggest that the parameters can also be estimated by various methods apart from curve-fitting.

Despite its attractive properties, this growth equation has rarely been used. The differential form of the equation [eqn (11)] is so simple that similar equations have occasionally been included in models (e.g. Aerts and van der Peijl, 1993). However, the value of this equation as a generic tool for empirical analysis and simulation has not previously been recognized. Hopefully this paper will encourage others to take advantage of this simple and powerful tool for the analysis and simulation of biological phenomena.

ACKNOWLEDGEMENTS

This work was funded by SOAEFD, Scotland. I thank David Hirst, David Causton, Philip Brain, Frank Moore and several anonymous referees for helpful and sometimes provocative comments and corrections. This manuscript was substantially revised while I was supported by a Centre of Excellence fellowship at Hokkaido University, Japan, funded by Monbusho.

LITERATURE CITED

- Aerts R, van der Peijl MJ. 1993. A simple model to explain the dominance of low-productive perennials in nutrient-poor habitats. *Oikos* **66**: 144–147.
- Alberda T, Sibma L. 1968. Dry matter production and light interception of crop surfaces. 3. Actual herbage production in different years as compared with potential values. *Journal of the British Grassland Society* **23**: 206–215.

- Birch CPD. 1995.** Modelling the vegetation dynamics of pasture including several species. In: Pollot GE, ed. *Grassland into the 21st century: challenges and opportunities. The British Grassland Society 50th anniversary conference*. Harrogate: British Grassland Society, 298–300.
- Birch CPD, Shaw MW. 1997.** When can reduced doses and pesticide mixtures delay the build-up of pesticide resistance? A mathematical model. *Journal of Applied Ecology* **34**: 1032–1042.
- Birch CPD, Werkman BR, Partridge LW. 1997.** A predictive model of vegetation dynamics under grazing. *18th International Grassland Congress*. Winnipeg and Saskatoon, Canada: 12:15–12:16.
- Brown RF, Mayer DG. 1988.** Representing cumulative germination. 2. The use of the Weibull function and other empirically derived curves. *Annals of Botany* **61**: 127–138.
- Causton DR, Venus JC. 1981.** *The biometry of plant growth*. London: Edward Arnold.
- France J, Thornley JHM. 1984.** *Mathematical models in agriculture*. London: Butterworths.
- Gaillard JM, Pontier D, Allaine D, Loison A, Herve JC, Heizmann A. 1997.** Variation in growth form and precocity at birth in eutherian mammals. *Proceedings of the Royal Society of London Series B—Biological Sciences* **264**: 859–868.
- Genstat 5 Committee. 1993.** *Genstat 5 release 3 reference manual*. Oxford: Clarendon Press.
- Grime JP, Hodgson JG, Hunt R. 1988.** *Comparative plant ecology. A functional approach to common British species*. London: Unwin Hyman.
- Huang SM, Titus SJ. 1994.** An age-independent individual tree height prediction model for boreal spruce aspen stands in Alberta. *Canadian Journal of Forest Research* **24**: 1295–1301.
- Huxley TH. 1870.** President's address to the Liverpool meeting of the British Association for the Advancement of Science. *Nature* **2**: 400–406.
- Meng FR, Meng CH, Tang SZ, Arp PA. 1997.** A new height growth model for dominant and codominant trees. *Forest Science* **43**: 348–354.
- Microsoft®. 1996.** *Excel 97*. Microsoft Corporation.
- Parresol BR. 1995.** Basal area growth for 15 tropical tree species in Puerto Rico. *Forest Ecology and Management* **73**: 211–219.
- Pearl E, Edwards TI, Miner JR. 1934.** The growth of *Cucumis melo* seedlings at different temperatures. *Journal of General Physiology* **17**: 687–700.
- Pienaar LV, Turnbull KJ. 1973.** The Chapman-Richards generalization of von Bertalanffy's growth model for basal area growth and yield in even-aged stands. *Forest Science* **19**: 2–22.
- Poorter H, Remkes C. 1990.** Leaf area ratio and net assimilation rate of 24 wild species differing in relative growth rate. *Oecologia* **83**: 553–559.
- Press WH, Teukolsky SA, Vetterling WT, Flannery BP. 1992.** *Numerical recipes in C. 2nd edn*. Cambridge: Cambridge University Press.
- Ratkowsky DA. 1983.** *Nonlinear regression modeling*. New York: Marcel Dekker Inc.
- Redfern D. 1993.** *The maple handbook: Maple V release 4. Waterpool Maple Inc*. New York: Springer-Verlag.
- Rennolls K. 1995.** Forest height growth modeling. *Forest Ecology and Management* **71**: 217–225.
- Richards FJ. 1959.** A flexible growth function for empirical use. *Journal of Experimental Botany* **10**: 290–300.
- Shvets V, Zeide B. 1996.** Investigating parameters of growth equations. *Canadian Journal of Forest Research* **26**: 1980–1990.
- Sigmaplot® 4.0. 1997.** *Transforms and regressions for Windows® 95, NT & 3.1 reference manual*. Chicago: SPSS Inc.
- Thornley JHM. 1995.** Shoot:root allocation with respect to C, N and P: an investigation and comparison of resistance and telenomic models. *Annals of Botany* **75**: 391–405.
- Thornley JHM, Johnson IR. 1990.** *Plant and crop modelling: a mathematical approach to plant and crop physiology*. Oxford: Oxford University Press.
- Verwijst T, Vonfircks HA. 1994.** Plant-response to temperature stress is characterized by an asymmetric sigmoid function. *Environmental and Experimental Botany* **34**: 69–74.
- Werker AR, Jaggard KW. 1997.** Modelling asymmetrical growth curves that rise and then fall: Applications to foliage dynamics of sugar beet (*Beta vulgaris* L.). *Annals of Botany* **79**: 657–665.
- Yang RC, Kozak A, Smith JHG. 1978.** The potential of Weibull-type functions as flexible growth curves. *Canadian Journal of Forest Research* **8**: 424–431.
- Zeide B. 1993.** Analysis of growth equations. *Forest Science* **39**: 594–616.

APPENDIX

Algebraic definition of the 11 curves used to generate the simulated dataset

All curves were constructed by joining pairs of logistic growth curves. A generic equation for a logistic growth curve is:

$$y = \frac{K-k}{1+e^{-a(t-t_0)}} + k \quad (\text{A } 1)$$

where k is the lower asymptote, K is the upper asymptote, a is a parameter of the initial acceleration of growth, equalling the maximum RGR when $k = 0$, and t_0 is the time at which $y = (K+k)/2$. This curve has no maxima or minima at finite times. It has a single point of inflexion at $t = t_0$.

(1) The curves used to generate the simulated dataset included four curves with the maximum rate of growth at size ≥ 50 .

$$y = \frac{100(2z-1)}{z(1+e^{-t})} \quad t \leq 0 \quad (\text{A } 2)$$

$$\Rightarrow \frac{dy}{dt} = \frac{100(2z-1)e^{-t}}{z(1+e^{-t})^2} \quad t \leq 0 \quad (\text{A } 3)$$

$$y = \frac{100(2z-1)}{2z-1+e^{-zt}} \quad t \geq 0 \quad (\text{A } 4)$$

$$\Rightarrow \frac{dy}{dt} = \frac{100(2z-1)ze^{-zt}}{(2z-1+e^{-zt})^2} \quad t \geq 0 \quad (\text{A } 5)$$

where $z = 1.0, 1.5, 2.0$ and 5.0 in each of the four curves.

(2) They also included seven curves with the maximum rate of growth at size < 50 .

$$y = \frac{100z}{z+(2-z)e^{-t}} \quad t \leq 0 \quad (\text{A } 6)$$

$$\Rightarrow \frac{dy}{dt} = \frac{100z(2-z)e^{-t}}{(z+(2-z)e^{-t})^2} \quad t \leq 0 \quad (\text{A } 7)$$

$$y = \frac{100(2-z)}{1+e^{-zt}} - 100(1-z) \quad t \geq 0 \quad (\text{A } 8)$$

$$\Rightarrow \frac{dy}{dt} = \frac{100z(2-z)e^{-zt}}{(1+e^{-zt})^2} \quad t \geq 0 \quad (\text{A } 9)$$

where $z = 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 0.75$ in each of the seven curves.

Substituting $t =$ large negative number in eqns (A 2) and (A 6) demonstrates that both sets of hybrid curves have lower asymptotes at $y = 0$. Substituting $t =$ large positive number in eqns (A 4) and (A 8) demonstrates that both sets have upper asymptotes at $y = 100$.

Substituting $t = 0$ in eqns (A 2) and (A 4) demonstrates that the upper and lower parts of the first set of hybrid curves meet at:

$$t = 0, \quad y = \frac{50(2z - 1)}{z}$$

Substituting $t = 0$ in eqns (A 3) and (A 5) demonstrates that the slopes of the upper and lower parts of the curves match at this point, with:

$$\frac{dy}{dt} = \frac{25(2z - 1)}{z}$$

Comparison with eqn (A 1) shows that the point of inflexion of eqn (A 2) is at $t = 0$, and of eqn (A 4) at $t \leq 0$, so the maximum slope of the hybrid curve is where the two logistic curves join.

Similarly, substituting $t = 0$ in eqns (A 6) and (A 8) demonstrates that the upper and lower parts of the second set of hybrid curves meet at:

$$t = 0, \quad y = 50z$$

Substituting $t = 0$ in eqns (A 7) and (A 9) demonstrates that the slopes of the upper and lower parts of the curves match at this point, with:

$$\frac{dy}{dt} = 25z(2 - z)$$

Comparison with eqn (A 1) shows that the point of inflexion of eqn (A 6) is at $t > 0$, and of eqn (A 8) at $t = 0$, so the maximum slope of the hybrid curve is where the two logistic curves join.

Thus these equations define hybrid curves that satisfy the following conditions: (1) they consist of two parts which join at $t = 0$; (2) they are continuous at $t = 0$; (3) they have continuous slopes at $t = 0$; (4) their maximum slopes are at $t = 0$; (5) their lower asymptotes are at $y = 0$; (6) their upper asymptotes are at $y = 100$.