Scientometrics, Vol. 25. No. 1 (1992) 5-46

CLASSIFICATION OF GROWTH MODELS BASED ON GROWTH RATES AND ITS APPLICATIONS

L. EGGHE*, I.K. RAVICHANDRA RAO**+

**LUC, Universitaire Campus, B-3590 Diepenbeek (Belgium)*
and
UIA, Universiteitsplein 1, B-2610 Wilrijk (Belgium)
***LUC, Universitaire Campus, B-3590 Diepenbeek (Belgium)*

(Received May 2, 1991)

In this paper, growth models are classified and characterised using two types of growth rates: from time t to t+1 and from time t to 2t. They are interesting in themselves but can also be used for a quick prediction of the type of growth model that is valid in a particular case. These ideas are applied on 20 data sets collected by *Wolfram*, *Chu* and *Lu*. We determine (using the above classification as well as via nonlinear regression techniques) that the power model (with exponent > 1) is the best growth model for Sci-Tech online databases, but that Gompertz-S-shaped distribution is the best for social sciences and humanities online databases.

I. Introduction

The 'law of exponential growth' is so well-known that it has become an everyday expression, used by virtually everybody. An important work is Ref. 3, in which (all over the book) growth aspects of literature and other phenomena are studied.

If C(t) denotes the number of items (e.g. articles in a database, books in a library, etc.) at time $t \geq 0$, then 'exponential growth' can be mathematically defined as

$$C(t) = C(O)\, e^{at}, \tag{1}$$

where $e^a > 1$, is the - in this case constant - growth rate from any time t to t+1:

$$C(t+1)/C(t) = e^a \tag{2}$$

*One of the authors (I.K.R. Rao) is grateful to the Belgian National Science Foundation (NFWO) for financial support in the period that he was a visiting professor in LUC.

+Permanent address: DRTC, 8th Mile, Mysore Rd, R.V. College P.O. Bangalore – 5600059, India.

Elsevier, Amsterdam – Oxford – New York – Tokyo
Akadémiai Kiadó

See Fig. 1 for the graph of Eq. (1).

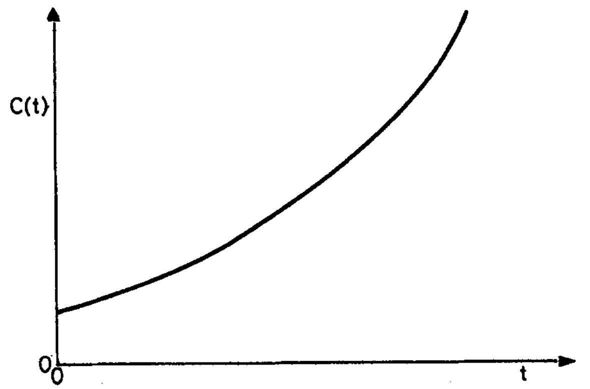

Fig. 1. Exponential growth

The number a is sometimes called the "Malthusian Parameter".[5] This number a originates from the differential equation

$$dC(t)/dt = a\,C(t) \qquad (3)$$

stating that the 'growth is proportional to the actual size', and a is exactly this factor of proportionality. We refer also to Refs 1 and 6 for some basic discussions on the exponential growth model.

However, already in Ref. 3, one finds lengthy discussions of ecountered deviations from exponential growth: at a certain time t_0, the growth declines giving rise to an S-shaped curve, as in Fig. 2.

Although many functions agree with the form of Fig. 2, one finds only reference to the so-called 'logistic curve'. Again, Ref. 3 is a basic reference.

The logistic function can be given mathematically as:[5]

$$C(t) = K/\{1+[K/C(0)-1]e^{-at}\} \qquad (4)$$

where $K > 0$ and $a > 0$. This function follows from the differential equation

$$dC(t)/dt = a(1-C(t)/K)C(t) \qquad (5)$$

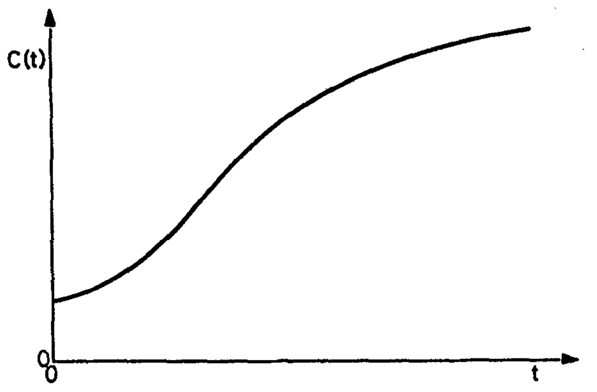

Fig. 2. S-shaped growth

to be compared with Eq. (3). Now the 'proportionality factor' a (1-C(t)/K) decreases when t increases, contrary to the case of Eq. (3). In case of logistic growth there is even an upper limit to C(t), since (4) has a horizontal asymptote (at height K) – see Fig. 3.

In this paper we will study another function with the same shape as in Fig. 3, namely Gompertz function[2]

$$C(t) = D.A^{B^t} \tag{6}$$

where $D > 0$, $\log A . \log B > 0$ (ensuring that C(t) increases).

If A, $B > 1$ then the function C has the form of Fig. 1: a steep increase (even steeper than the exponential function). If $0 < A, B < 1$, then the shape of C is as in Fig. 3. In the sequel we will show that this last case is a better model than the logistic one (in cases of S-shaped growth). As far as we know, it is the first time that Gompertz function is studied in informetrics and this paper will show that it is an important model.

The exponential function is convex while the logistic (or Gompertz with $0<A$, $B<1$) has a convex part and then (for larger t) becomes concave.

In one paper in informetrics, we encountered a completely concave function: the model of *Ware*.[7] The mathematical function is:

$$C(t) = \delta(1-\varphi^{-t}) \tag{7}$$

where $\delta > 0$ and $\varphi > 1$ and has the form as in Fig. 4.

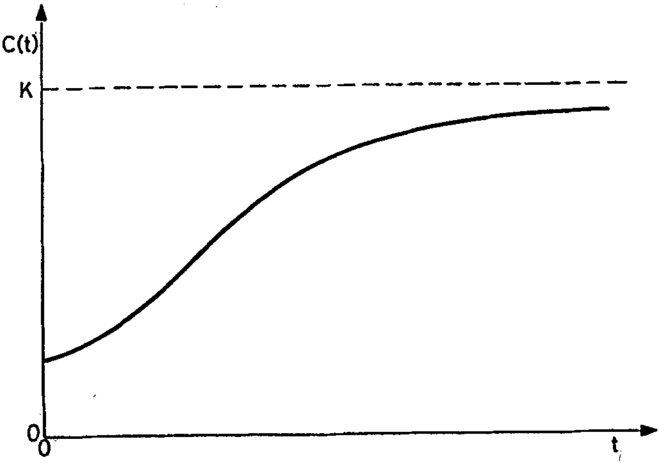

Fig. 3. Logistic growth

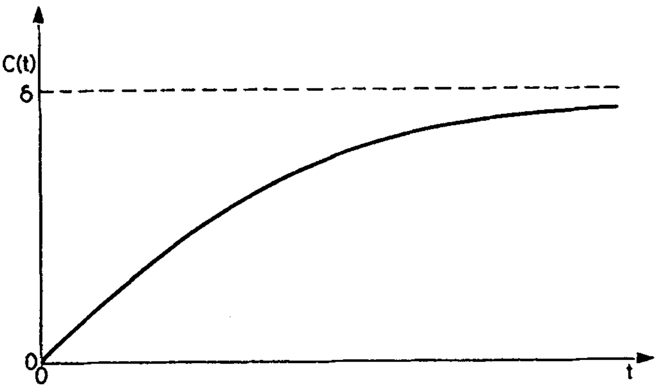

Fig. 4. Ware's model

Our calculations will however show that Ware's function is not applicable to the data that we will investigate.

Finally in Ref. 8 we encountered the power model of growth:

$$C(t) = a + \beta t^{\gamma} \tag{8}$$

where α, $\beta > 0$. For $0 < \gamma < 1$, this function is concave but without an upper limit (as in Fig. 4) – see Fig. 5a. For $\gamma = 1$, we have linear growth (evidently) – see Fig. 5b. For $\gamma > 1$, we have convex growth – see Fig. 5c.

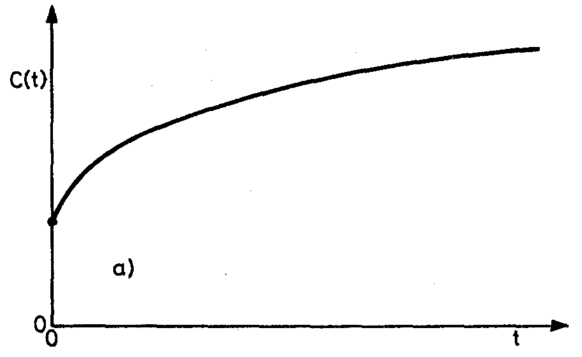

Fig. 5a. Power model ($0 < \gamma < 1$)

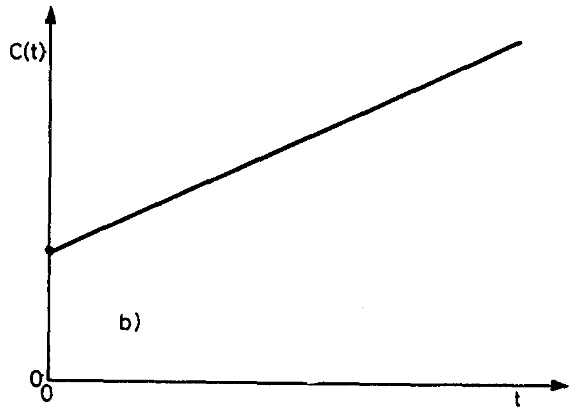

Fig. 5b. Linear model ($\gamma = 1$)

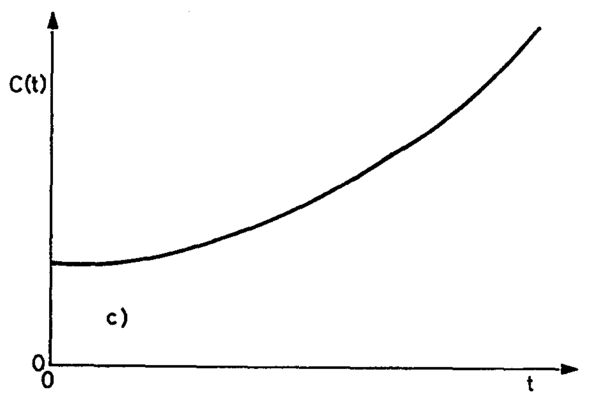

Fig. 5c. Power model ($\gamma > 1$)

In our study, we have used the data (20 data sets describing the content of 20 online datebases between 1965 and 1987) collected by *Wolfram* et al.[8]

But in addition to only applying some 'best fitting techniques' in statistics, we give our study also a qualitative dimension. Indeed, growth tendencies are not easy to distinguish from each other. One evident reason is that all the data are increasing (certainly in the case of cumulative data sets as e.g. the yearly content of online databases). Convexity, concavity (or a mixture) can be seen when graphing the growth data but even then we have several distinct models e.g.: all Gompertz curves (A, $B > 1$), exponential curves and power curves ($\gamma > 1$) are convexly increasing! Also, all power models ($0 < \gamma < 1$) and Ware's function are concavely increasing!

Inspired by our obsolescence study,[4] we will investigate the growth rate of the data. The most evidental rate to look at is the function

$$\alpha_1(t) = C(t+1)/C(t) \tag{9}$$

i.e. the rate of growth from year t to year $t + 1$. However, in our study the following growth rate function is also very important:

$$\alpha_2(t) = C(2t)/C(t) \tag{10}$$

The function (9) was already studied in Ref. 4 with successful applications. In this paper however, function (10) is also very important in the sense that – only on the basis of α_2 – we will be able to predict the exact model C for the growth in most cases of the 20 datasets. In order to obtain a complete classification of growth models C in terms of growth rates functions (α_1 and α_2) we will, however, study both α_1 and α_2 in the next section.

The third section is then devoted to the statistical study of the 20 data sets of Ref. 8. We apply the method of non-linear regression, by using the STATGRAPHICS – statistical package. Our results are surprising in the sense that

- the exponential model never occurs;
- only the power models ($\gamma > 1$) or Gompertz models ($\log A \log B > 0$, $0 < A, B < 1$) are applicable;
- the power models are best for Sci Tech online databases, while the Gompertz models are best for the Social Sciences and Humanities online databases (indicating a faster 'growth rate' for science and technology than for social sciences and humanities).

The fourth section shows that the best statistical models (found in the third section via the statistical methods) could have been predicted based on our classification of Section 2 (mainly based on visual inspection of the table of the function α_2), hence showing the value of the method. So this method constitutes a very easy and mathematically founded way of direct recognition of a growth model. Furthermore, even when a statistical fitting method has been applied, it helps to recognise and explain 'the right model'.

The fifth section discusses some 'difference aspects' of the models that have been discussed here and the sixth section repeats the main conclusions.

II. Classification of growth models based on growth rate functions

II. 1. The growth rate functions α_1 *and* α_2

Let C(t) denote the growth function (theoretical or concrete data) for t = 1, 2, ... In the same way as we defined the aging rate in Ref. 4, we can now define:

$$\alpha_1(t) = C(t+1)/C(t) \tag{9}$$

for t = 1, 2,.... . In the course of our research for this paper it became apparent that also another growth rate function is needed. We define:

$$\alpha_2(t) = C(2t)/C(t) \tag{10}$$

for t = 1, 2,.... . Note that, if there are in total N observations (i.e. t = 0, 1,..., N-1) we have N-1 values of α_1 and N/2 values of α_2. This is a small disadvantage of α_2 but in most cases this is not a serious drawback: in our test cases N = 20 and hence α_2 has 10 values, enough to see the graph of α_2. The function α_2 is very natural to study: it compares the growth after a double time period (from t to 2t). α_1 is called the *first growth rate* function and α_2 is called the *second growth rate function*. They are the basic tools in this paper. The base idea is that the graphs of α_1 and α_2 are much more different than their corresponding graphs of the different growth models C (this will be shown in this section).
We note the following theoretical relation between α_1 and α_2:

$$\alpha_2(t) = C(2t)/C(t) = (C(2t)/C(2t\text{-}1))\ (C(2t\text{-}1)/C(2t\text{-}2))$$
$$\dots (C(t+1)/C(t))$$

$$\alpha_2(t) = \alpha_1(2t\text{-}1)\ \alpha_1(2t\text{-}2) \dots \alpha_1(t) \tag{11}$$

We will now investigate what the graphs are of α_1 and α_2 for the following growth models C: exponential, logistic, power, Ware and Gompertz. At the same time we will investigate the graph of C itself, for the sake of completeness. In the sequel, t will be a continuous variable: $t \geq 0$.

II. 2. The exponential model

We rewrite function (1) into

$$C(t) = c.g^t \tag{12}$$

where $c > 0$, $g > 1$, $t \geq 0$. The graph is as in Fig. 1.
Evidently

$$\alpha_1(t) = g \tag{13}$$

a constant function, above 1 (see Fig. 6a).

Furthermore,

$$\alpha_2(t) = g^t \tag{14}$$

for all $t \geq 0$ and hence α_2 looks as in Fig. 6b.

II. 3. The logistic model

To simplify the notation, Eq. (4) is rewritten as

$$C(t) = 1/(k+ab^t) \tag{15}$$

where $k, a > 0$, $0 < b < 1$, $t \geq 0$. The graph is as in Fig. 3 ($K = 1/k$). Now

$$\alpha_1(t) = (k+ab^t)/(k+ab^{t+1}) \tag{16}$$

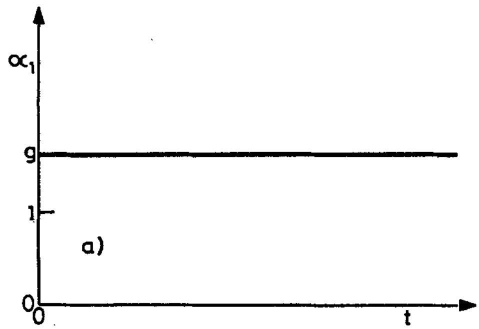

Fig. 6a. α_1 for the exponential model

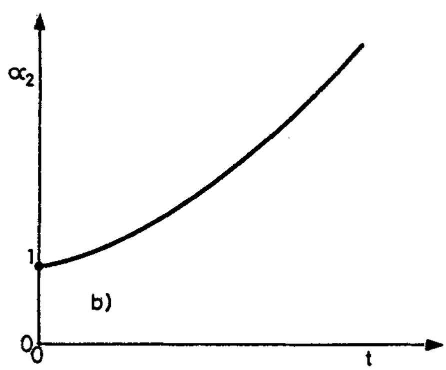

Fig. 6b. α_2 for the exponential model

We have (we omit the easy proof): $\lim \alpha_1(t) = 1$, $\alpha_1(0) = (k+a)/(k+ab) > 1$, $\alpha_1'(t) < 0$ for all $t \geq 0$ and $\alpha_1'(0) > -\infty$.
Hence α_1 has a graph as in Fig. 7a.

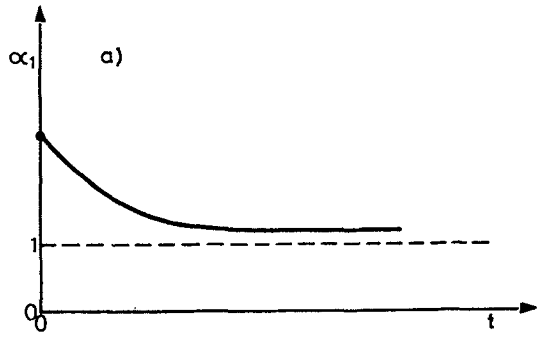

Fig. 7a. α_1 for the logistic model

$$\alpha_2(t) = (k+ab^t)/(k+ab^{2t}) \tag{17}$$

Now $\lim_{t\to\infty} \alpha_2(t) = 1$, $\alpha_2(0) = 1$. Furthermore,

$$\alpha_2'(t) = a(\log b)b^t\,(k\text{-}2\,kb^t\text{-}ab^{2t})/(k+ab^{2t})^2.$$

Hence, $\alpha_2'(t) < 0$, for $t > t_0$ and $\alpha_2'(t) > 0$ for $t < t_0$ where

$$b^{t_0} = (-k + \sqrt{k^2+ak})/a > 0$$

See Fig. 7b for the graph of α_2.

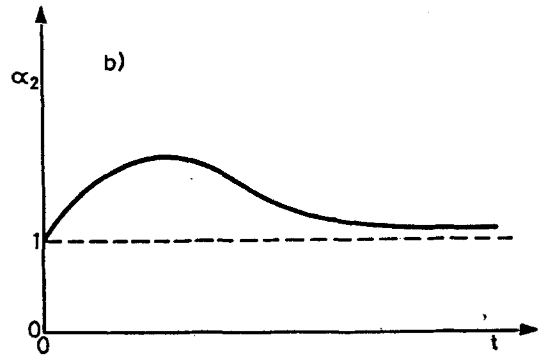

Fig. 7b. α_2 for the logistic model

II. 4. Gompertz model

A well-known growth model (but until now absent in informetric studies) is the one of Gompertz:

$$C(t) = D.A^{B^t} \tag{18}$$

where $D > 0$, $t \geq 0$. Furthermore, only if $\log A \log B > 0$, (18) represents an increasing function. As shown in Ref. 2, we have graph 8a in case $0 < A, B < 1$ and graph 8b in case $A, B > 1$. Since, if $A, B > 1$, function (18) represents a 'super-increase' (much faster than the exponential) and since it is well known that exponential increase is often too much,[3] we conjecture already here that Fig. 8b will not occur often (this will be seen in the sequel).

$$\alpha_1(t) = A^{B^{t+1}-B^t} \tag{19}$$

Hence

$$\alpha_1'(t) = A^{B^{t+1}-B^t}(\log A \log B)(B-1)B^t,$$

concluding that $\alpha_1' > 0$ for $B > 1$ and $\alpha_1' < 0$ for $0 < B < 1$.

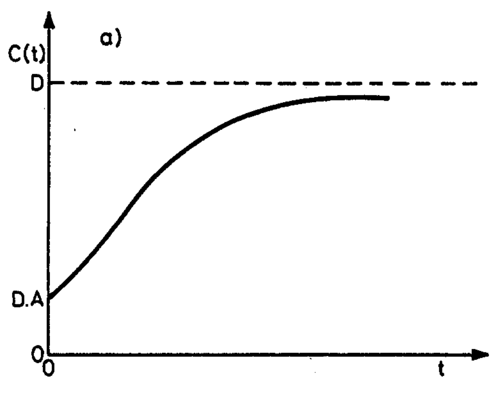

Fig. 8a. Gompertz model for $\log A \log B > 0$, $0 < A, B < 1$

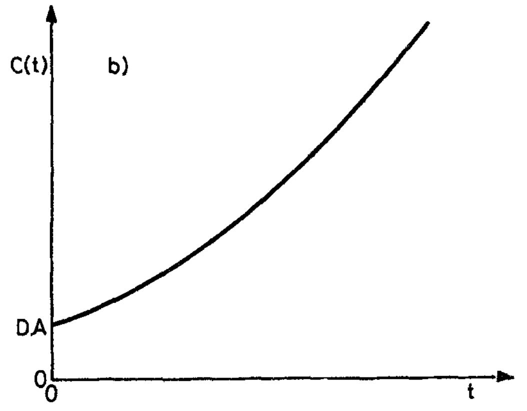

Fig. 8b. Gompertz model for log A log B > 0, A, B > 1

Furthermore,

$$\alpha_1(0) = A^{B-1} > 1$$

and

$$\lim \alpha_1(t) \begin{cases} = 1 \text{ for } 0 < B < 1 \\ = +\infty \text{ for } B > 1. \end{cases}$$

This gives Fig. 8c for the graphs of α_1.

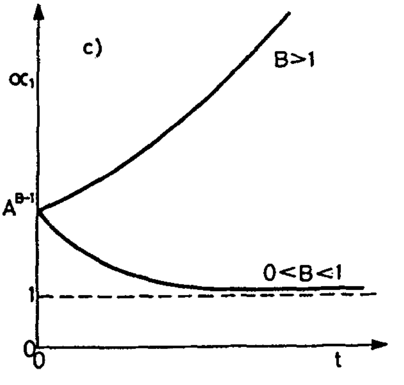

Fig. 8c. α_1 for Gompertz models

$$\alpha_2(t) = A^{B^{2t}-B^t} \tag{20}$$

We have $\alpha_2(0) = 1$,

$$\lim_{t\to\infty} \alpha_2(t) \begin{cases} = 1 & \text{for } 0 < B < 1 \\ = +\infty & \text{for } B > 1 \end{cases}$$

$$\alpha_2'(t) = A^{B^{2t}-B^t} \log A \log B\,(2B^{2t}-B^t)$$

Hence $\alpha_2'(t_0)$ can only be zero for a $t_0 > 0$ if $0 < B < 1$. In this case $\alpha_2'(t) > 0$ (for $t < t_0$) and $\alpha_2'(t) < 0$ (for $t > t_0$). If $B > 1$ then $\alpha_2'(t) > 0$ for every $t \geq 0$. Furthermore

$$\lim_{\substack{t\to 0\\ >}} \alpha_2'(t) = \log A \log B > 0.$$

Therefore, the graphs of α_2 are as in Fig. 8d.

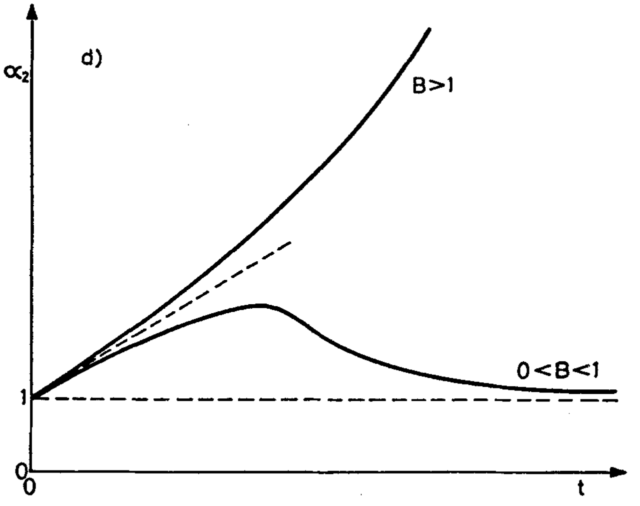

Fig. 8d. α_2 for Gompertz models

II. 5. Ware's model

Here C is the function:

$$C(t) = \delta(1-\varphi^{-t}) \tag{21}$$

where $\delta > 0$, $\varphi > 1$, $t \geq 0$. The graph is as in Fig. 4.

$$\alpha_1(t) = (1-\varphi^{-t-1})/(1-\varphi^{-t}) \tag{22}$$

Hence $\alpha_1'(t) < 0$ for all $t \geq 0$, $\lim \alpha_1(t) = 1$ and
$\alpha_1(0+) = \lim\limits_{t \underset{>}{\to} 0} \alpha_1(t) = +\infty$. We have Fig. 9a.

$$\alpha_2(t) = (1 - \varphi^{-2t})/(1 - \varphi^{-t}) \tag{23}$$

Hence

$$\alpha_2'(t) = [-(\varphi^{t}-1)^2 \log \varphi]/[\varphi^{3t}(1-\varphi^{-t})^2] < 0$$

for all $t \geq 0$. Also $\alpha_2(0+) = 2$, $\lim\limits_{t\to\infty} \alpha_2(t) = 1$ and

$\alpha_2'(0+) = -\log \varphi < 0$. Hence we have the graph of Fig. 9b.

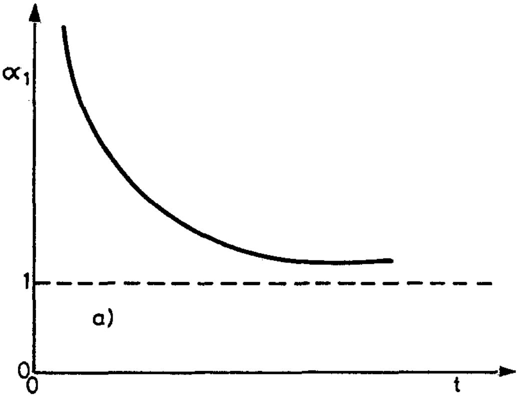

Fig. 9a. α_1 for Ware's model

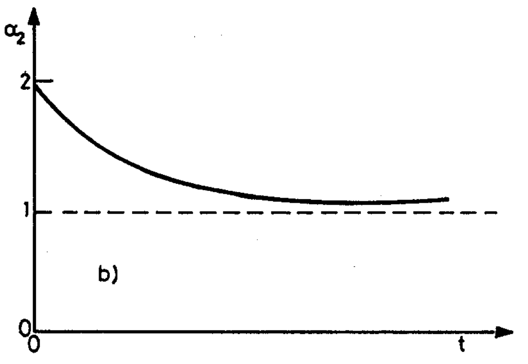

Fig. 9b. α_2 for Ware's model

Finally, we study the different cases of the power law:

II. 6. The power model

Here we study the functions

$$C(t) = \alpha + \beta t^{\gamma}, \tag{24}$$

where γ, $\beta > 0$ and $\alpha \geq 0$ (we suppose $C(0) \geq 0$ indeed, although some statistical fittings exceptionally give an $\alpha < 0$, we exclude this case from our mathematical models).

If $\gamma = 1$, then model (24) is called the linear law, evidently. The graphs are as in Figs 5a, b, c,. Now

$$\alpha_1(t) = [\alpha + \beta(t+1)^{\gamma}]/(\alpha + \beta t^{\gamma}) \tag{25}$$

Hence, $\alpha_1(0) = (\alpha + \beta)/\alpha > 1, \lim_{t\to\infty} \alpha_1(t) = 1$ and

$$\alpha_1'(t) = (\alpha\beta\gamma[(t+1)^{\gamma-1} - t^{\gamma-1}] - \beta^2\gamma t^{\gamma-1}(t+1)^{\gamma-1})/(\alpha + \beta t^{\gamma})^2$$

Consequently, if $\gamma > 1$ and $\alpha > 0$, there is a $t_0 > 0$ such that $\alpha_1'(t) > 0$ for $t < t_0$ and $\alpha_1'(t) < 0$ for $t > t_0$ and $\alpha_1'(t_0) = 0$. Hence the graph is as in Fig. 10a (for $\alpha \neq 0$). If $\alpha = 0$, then $\alpha_1(0+) = +\infty$ and $\alpha_1'(t) < 0$ for every $t \geq 0$. Hence the graph is as in Fig. 10b ($\alpha = 0$). For $0 < \gamma \leq 1$ we have that $\alpha_1'(t) < 0$ for every $t \geq 0$. For $\alpha > 0$, see Fig. 10c and for $\alpha = 0$, again Fig. 10b can be used. These last two figures hence also contain the linear case ($\gamma = 1$).

$$\alpha_2(t) = (\alpha+\beta(2t)^\gamma)/(\alpha+\beta t^\gamma) \tag{26}$$

Now $\alpha_2(0) = 1$, $\lim_{t\to\infty} \alpha_2(t) = 2^\gamma > 1$,

$$\alpha_2'(t) = [\alpha\beta\gamma t^{\gamma-1}(2^\gamma-1)]/(\alpha+\beta t^\gamma)^2$$

Hence $\alpha_2'(t) > 0$ for $\alpha > 0$ and $\alpha_2'(t) = 0$ for $\alpha = 0$. Also

$$\alpha_2'(0+) \begin{cases} = 0 & \text{for } \gamma > 1 \\ = +\infty & \text{for } 0 < \gamma < 1 \\ = \beta > 0 & \text{for } \gamma = 1 \end{cases}$$

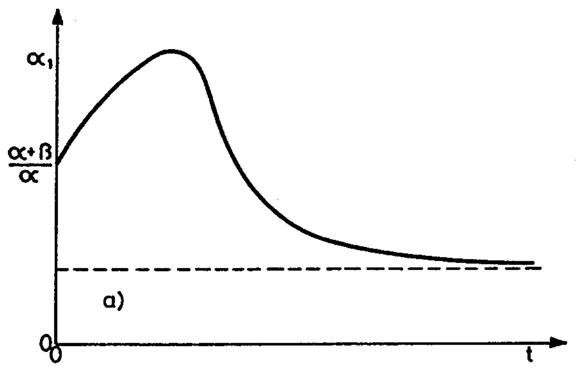

Fig. 10a. α_1 for the power model ($\gamma > 1$, $\alpha > 0$)

Hence, we have the following graphs of α_2:
Fig. 10d for $0 < \gamma < 1$, $\alpha > 0$, Fig. 10e for $\gamma = 1$, $\alpha > 0$, Fig. 10f for $\gamma > 1$, $\alpha > 0$, Fig. 10g for $\gamma > 0$, $\alpha = 0$.

Note:
The point $t_0 > 0$ in which $\alpha_1'(t_0) = 0$ in Fig. 10a ($\gamma > 1$, $\alpha > 0$) can be very small (if γ is close to 1). For $\gamma \leq 1$ we even have that t_0 does not exist anymore (cf. Fig. 10b and c). In these cases, Fig. 10a can be regarded as a decreasing graph. These cases are most occurring since growths with high γ-powers are not likely to happen.

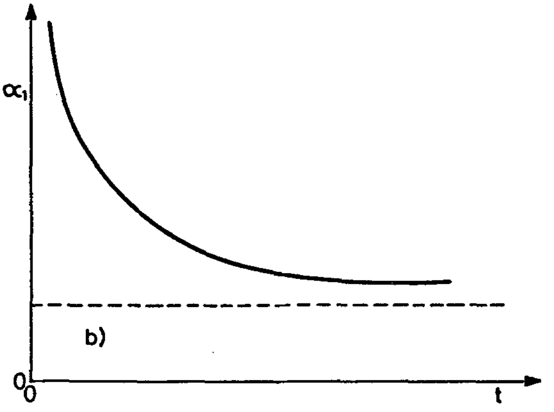

Fig. 10b. α_1 for the power model ($\alpha = 0, \gamma > 0$)

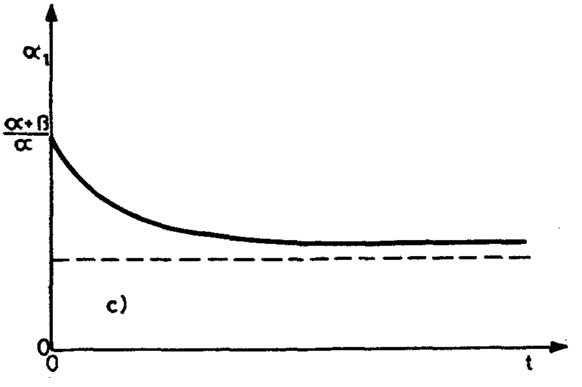

Fig. 10c. α_1 for the power model ($0 < \gamma \leq 1, \alpha > 0$)

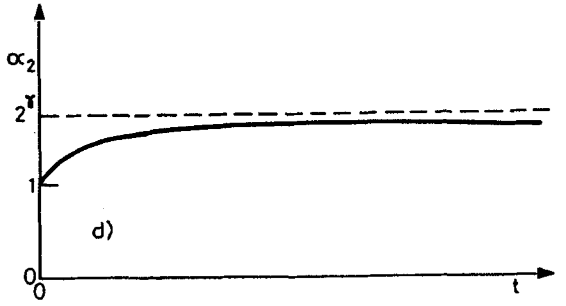

Fig. 10d. α_2 for the power model ($0 < \gamma < 1, \alpha > 0$)

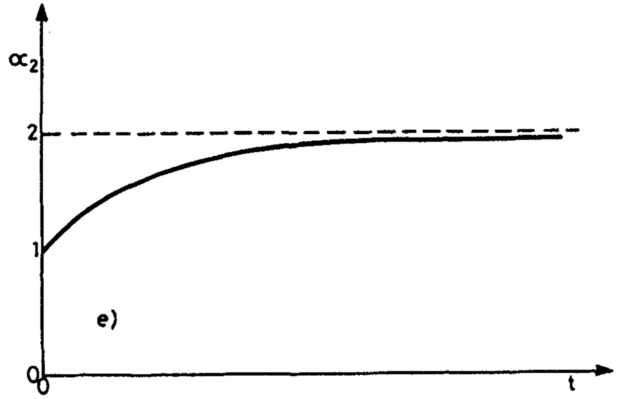

Fig. 10e. α_2 for the power model ($\gamma = 1, \alpha > 0$)

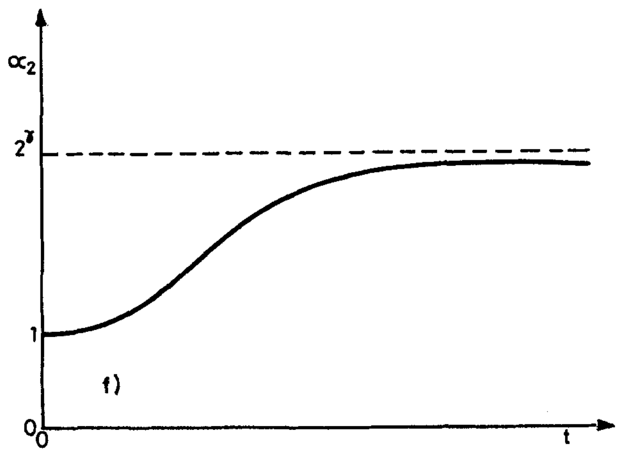

Fig. 10f. α_2 for the power model ($\gamma > 1, \alpha > 0$)

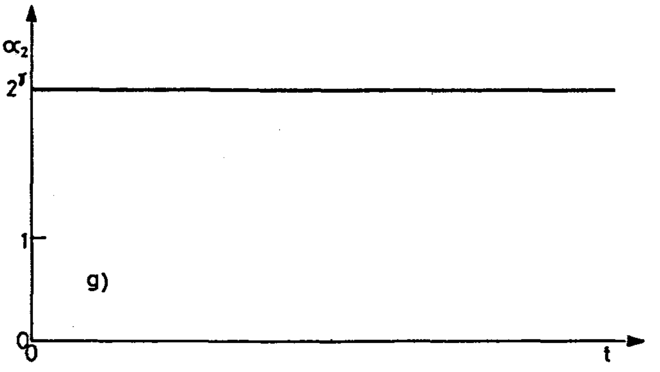

Fig. 10g. α_2 for the power model ($\gamma > 0$, $\alpha = 0$)

II. 7. Simple classification

We make the following simple classification, which can be used without requiring any calculations or statistical fits:

Type:	Function:
I	increasing
II	constant
III	decreasing
IV	increasing and then decreasing

The study in this section allows for the classification shown in Table 1.

Table 1
Classification of growth models based on growth rate functions

(α^1,α^2)	C-function
(*) (II, I)	exponential
(III, IV)	logistic or Gompertz ($0 < A, B < 1$)
(*) (I, I)	Gompertz ($A, B > 1$)
(III, I)	Power ($\alpha > 0, 0 < \gamma \leq 1$)
(*) (IV, I)	Power ($\alpha > 0, \gamma > 1$)
(III, II)	Power ($\alpha = 0$)
(III, III)	Ware

Note that in the cases marked with (*) the calculation of α_2 is not even needed. Only if α_1 is of type III, then α_2 must be determined. It must however be said that in all cases that we encountered, α_1 was of type III (decreasing).

Note also that only in case (III, IV) we cannot distinguish between the logistic and Gompertz ($0 < A, B < 1$) model. It must however be said that both models look much alike (growth model C as well as growth rate functions α_1, α_2). For a distinction between the two models a statistical fit is needed (see further).

So this method of determining growth models goes back to the intrinsic growth rate properties of the data and allows for a good understanding of what is really going on. Our advise is, however, to combine it with a statistical fitting procedure so that conclusions are based on two different methods. But for a quick and yet mathematically sound idea of what type of growth model one has, the above classification can help.

The next section is devoted to the statistical fittings of the 20 datasets in Ref. 8. Then, in the following section, a comparison of both methods will be given.

III. Statistical fittings

We used the 20 data sets as compiled by *Wolfram*, *Chu* and *Lu*,[8] which we give here for the sake of completeness. The data sets refer to the following online databases (taken from Ref. 8) (see Table 2).

We used the STATGRAPHICS (version 2.0) package and more particularly, the nonlinear regression section. We obtained the following fits for the respective models: see Tables 4a, b, c, d for the models: Gompertz, power, logistic, exponential. We excluded Ware's model here since this only allows for a concave function, while all our data are convex or S-shaped: it is common sense not to try to fit a convex or S-shaped function by a concave one!

For reasons of general interest we include also the parameters of the linear fit ($\gamma = 1$) ($C(t) = \alpha + \beta t$) and of the multiplicative fit ($\alpha = 0$) ($C(t) = \beta t^\gamma$). They are given in Tables 4e and f.

Table 2
Databases used

Data set	Subject area	Database name
1	Humanities	PHILOSOPHER'S INDEX
2	Humanities	RELIGION INDEX
3	Social Sciences	ECONOMIC LITERATURE INDEX
4	Social Sciences	ERIC
5	Social Sciences	EXCEPTIONAL CHILD EDUCATION RESOURCES
6	Social Sciences	LIBRARY AND INFORMATION SCIENCE ABSTRACTS
7	Social Sciences	SOCIOLOGICAL ABSTRACTS
8	Sci. & Techn.	BIOSIS PREVIEWS
9	Sci. & Techn.	CA SEARCH
10	Sci. & Techn.	FOOD SCIENCE AND TECHNOLOGY ABS.
11	Sci. & Techn.	GEOREF
12	Sci. & Techn.	INSPEC
13	Sci. & Techn.	MEDLINE
14	Sci. & Techn.	MENTAL HEALTH ABSTRACTS
15	Sci. & Techn.	METADEX
16	Sci. & Techn.	NTIS
17	Sci. & Techn.	OCEANIC ABSTRACTS
18	Sci. & Techn.	PAPERCHEM
19	Sci. & Techn.	SMOKING AND HEALTH
20	Sci. & Techn.	WORLD ALUMINUM ABSTRACTS

Table 3
Cumulated data sets used

Year	Set 1	Set 2	Set 3	Set 4	Set 5
1968	3106	1862	3	8697	933
1969	6669	3343	4472	19011	2371
1970	10613	8818	9507	41889	4581
1971	15030	13634	14442	73718	7161
1972	18966	19290	20101	106721	9976
1973	23541	25050	26033	139705	12829
1974	28476	32112	31965	171428	16300
1975	33520	47921	37944	205178	19871
1976	38451	64367	44315	237937	24338
1977	43100	80920	51332	269635	28913
1978	48218	96907	58929	301681	33555
1979	54179	113096	70649	325516	37914
1980	60176	132280	83182	347930	41889
1981	66616	151927	96788	368534	44845
1982	73440	174790	105087	389428	47637
1983	79951	200785	119069	410733	50408
1984	86521	229603	127809	432919	53235
1985	93001	257350	137212	454208	56031
1986	100021	277050	146516	474840	58747
1987	106275	282833	155108	493815	61147

Year	Set 6	Set 7	Set 8	Set 9	Set 10
1968	736	6144	158642	239464	7578
1969	3400	12292	362954	510340	25148
1970	5940	19436	574713	803986	44716
1971	8516	26279	808829	1123763	61826
1972	11501	32958	1052663	1460691	78822
1973	14479	38713	1292709	1818772	97327
1974	17174	44483	1518867	2192050	117353
1975	20855	51007	1765408	2566744	136203
1976	24645	57825	2018140	2961158	154424
1977	28816	67270	2297657	3377268	173392
1978	33322	79359	2574759	3799002	194629
1979	38710	90205	2844465	4227927	215017
1980	44179	100361	3147250	4668272	235254
1981	49830	109719	3493858	5114772	254764
1982	55793	119290	3840122	5563322	274444
1983	61779	128176	4115014	6016913	293823
1984	67678	137665	4485854	6476704	313176
1985	72665	147915	4918660	6939190	332119
1986	78197	161360	5380812	7419037	348571
1987	82029	170667	5873950	7875024	361886

(Table 3 cont.)

Year	Set 11	Set 12	Set 13	Set 14	Set 15
1968	32963	47955	200831	22276	24950
1969	71754	136912	408652	43654	51213
1970	116480	239043	617540	69948	78122
1971	159343	357646	832257	96896	105442
1972	202817	487550	1051246	128043	131665
1973	249091	620755	1274979	157116	159748
1974	295854	751123	1504333	188107	188936
1975	342055	891639	1747810	218199	219302
1976	386857	1033184	1993887	251931	250934
1977	432506	1180878	2241856	287790	284422
1978	487614	1343073	2495629	318286	318478
1979	545337	1505179	2757284	340631	354192
1980	604406	1677420	3018488	363833	391084
1981	667933	1856777	3282016	394927	429309
1982	732285	2025882	3555349	415192	468218
1983	790857	2231925	3840659	427177	506053
1984	853469	2448392	4135664	432164	547463
1985	910903	2663299	4439808	435012	588350
1986	962784	2882102	4754708	436410	628473
1987	1004748	3115826	5076104	437869	661729

Year	Set 16	Set 17	Set 18	Set 19	Set 20
1968	42103	8689	5660	1562	4263
1969	85665	19788	12030	3009	8905
1970	133356	29334	18638	4376	14374
1971	184486	38538	25641	5463	19716
1972	241043	46772	32232	6890	25554
1973	298574	52769	38884	8315	31761
1974	352701	56605	45276	9796	38206
1975	404369	63485	52477	11168	45072
1976	455323	68620	59486	12732	51769
1977	510871	73482	66409	14355	57267
1978	569296	80264	73657	15937	62117
1979	627194	87877	80962	17552	67397
1980	688547	94578	88264	19026	72756
1981	752186	104088	96069	20649	78504
1982	811190	113369	103640	22375	83959
1983	868691	123388	111524	24552	89444
1984	925320	131697	118881	26524	95070
1085	980161	141671	126422	28330	100657
1986	1035290	149860	134194	30038	106943
1987	1083528	155256	141180	31058	112135

Table 4a
Parameters of the Gompertz model $C(t) = D.A^{B^t}$

Data set	D	A	B
1	198017.107	0.033	0.915
2	576787.486	0.005	0.897
3	285913.933	0.015	0.903
4	548169.123	0.037	0.845
5	73182.7042	0.214	0.8538
6	159668.622	0.018	0.909
7	352803.012	0.032	0.921
8	13287123.8	0.0*	0.9
9	12732165.5	0.0*	0.9
10	491968.684	0.048	0.887
11	1673167.02	0.04	0.91
12	6232544.30	0.03	0.92
13	8568176.78	0.05	0.91
14	486430.701	0.05	0.822
15	1167507.00	0.04	0.91
16	1552913.15	0.05	0.90
17	307433.016	0.066	0.930
18	211350.825	0.053	0.903
19	52874.9863	0.0500	0.9127
20	145034.889	0.057	0.885

*If any of the parameters require 8 digits to display/print the integer part of the values, we have observed that in the STAGRAPHICS package for all the parameters only one digit is being displayed/printed in the decimal position. Thus it printed only the first digit of the decimal portion; in this case it is zero.

Table 4b
Parameters for the power model $C(t) = \alpha + \beta t^{\gamma}$

Data set	α	β	γ
1	4509.20008	2355.34772	1.27897
2	-186.36009	1681.99903	1.76064
3	2005.67066	2165.68835	1.45650
4	-15706.6931	40721.7780	0.8674
5	-1355.56898	3049.3902	1.04002
6	1952.53386	1169.66358	1.44458
7	9545.02345	3212.88146	1.33653
8	299664.0	104367.0	1.340
9	270254.941	225464.527	1.195
10	6699.051	17911.6013	1.0212
11	42779.9292	29566.2179	1.1888
12	94255.5284	56396.8639	1.3471
13	257981.760	150898.222	1.173
14	-5166.5136	53851.3513	0.7506
15	33509.8252	17304.9702	1.2216
16	38786.0597	46419.7381	1.0616
17	15437.87900	5668.39620	1.09250
18	6386.08829	5679.84153	1.07621
19	1935.20341	969.53587	1.16090
20	2812.42877	6252.94242	0.97233

Table 4c
Parameters of the logistic model $C(t) = 1/(k+ab^t)$

Data set	k	a	b
1	0.00000755	0.00010631	0.81254766
2	0.00000285	0.00013012	0.75844254
3	0.00000534	0.00013110	0.77960827
4	0.00000204	0.00002480	0.74718214
5	0.00001592	0.00028203	0.74361837
6	0.00000979	0.00021068	0.79105917
7	0.00000452	0.00006359	0.81645410
8	0.00000011	0.00000188	0.82983662
9	0.00000011	0.00000142	0.80414483
10	0.00000250	0.00002768	0.79135401
11	0.00000083	0.00001031	0.80797878
12	0.00000025	0.00000421	0.81157002
13	0.00000016	0.00000183	0.81859200
14	0.00000221	0.00002193	0.72523916
15	0.00000123	0.00001514	0.81510474
16	0.00000082	0.00000860	0.80152600
17	0.00000475	0.00003837	0.84930120
18	0.00000613	0.00006209	0.81053107
19	0.00002621	0.00027663	0.81917840
20	0.00000827	0.00007919	0.79542581

Table 4d
Parameters for the exponential model $C(t) = cg^t$

Data set	c	g
1	8315.9447	1.1663
2	5773.536	1.2710
3	2022.4805	1.3300
4	36930.757	1.1828
5	3651.5244	1.1974
6	3830.0763	1.2074
7	14657.977	1.1590
8	445298.45	1.1654
9	643321.58	1.1646
10	32006.683	1.1644
11	89922.188	1.1589
12	179261.34	1.1891
13	494399.91	1.1503
14	61938.671	1.1392
15	61322.371	1.1543
16	110172.21	1.1510
17	22502.912	1.1223
18	14968.301	1.1468
19	3466.1855	1.1418
20	11876.418	1.1491

Table 4e
Parameters for the linear model $C(t) = \alpha + \beta t$

Data set	α	β
1	-2541.4	5477.36
2	-41401.8	16010.4
3	-13818.1	8509.6
4	7957.17	26917.8
5	-2126.44	3448.47
6	-6151.33	4438.27
7	-3661.93	8865.07
8	-137718.0	290946.0
9	-121908.0	408382.0
10	4396.1	19118.7
11	-6178.99	52492.8
12	-1507.59	160588.0
13	37601.7	255142.0
14	41178.6	24431.0
15	-2679.87	33903.6
16	19833.2	56069.6
17	11773.0	7393.0
18	3421.11	7174.23
19	648.6	1582.82
20	3758.64	5740.51

Table 4f
Parameters for the multiplicative model $C(t) = \beta t^{\gamma}$

Data set	β	γ
1	7.97343	1.18745
2	7.07962	1.82329
3	4.84574	2.58664
4	9.19993	1.37567
5	6.87527	1.43568
6	6.89187	1.48791
7	8.61372	1.12454
8	11.9532	1.18457
9	12.3213	1.18154
10	9.22519	1.22553
11	10.3838	1.145
12	10.8611	1.36113
13	12.1519	1.0817
14	10.0493	1.05002
15	10.0507	1.10388
16	10.6194	1.09916
17	9.20629	0.902816
18	8.65021	1.06979
19	7.26148	1.0151
20	8.3774	1.09854

The quality of the fits is shown in Table 5a (giving R^2-values) and Table 5b (giving the residual standard deviation). A(*) means that this model is the best. There is a complete agreement between both tables.

For reasons of general interest we also included the quality values for the linear and the multiplicative models. Of course, the general power law, having one more free parameter, gives always closer fits.

We think that this study reveals a few remarkable conclusions:

1. The exponential model nor the logistic model (except in one case: set 14) is the best model.
2. In case of an S-shaped growth curve, the Gompertz model fits best (except in one case: set 14, but even then, Gompertz is very close to the logistic model).
3. In all the other cases, the power model fits best (mainly all with $\gamma > 1$, i.e. convex growth).

Table 5a

R-square : the measure of proportion of total variation about the mean explained by the regression

Data Sets	Gompertz model	Power model general	Power model linear (γ = 1 fixed)	Power model multiplicative (α = 0 fixed)	Logistic model	Exponential model
1	0.998394	0.999581(*)	0.9914	0.9989	0.995504	0.8880
2	0.998094 (*)	0.996614	0.9494	0.9894	0.997423	0.9069
3	0.997939 (*)	0.996971	0.9772	0.9772	0.997337	0.5022
4	0.99718 (*)	0.944453	0.9913	0.9822	0.989995	0.7754
5	0.999431 (*)	0.993017	0.9928	0.9958	0.996707	0.8308
6	0.999068 (*)	0.998608	0.9798	0.9935	0.997928	0.8450
7	0.996457	0.996741 (*)	0.9895	0.9945	0.994479	0.9068
8	0.996388	0.997906 (*)	0.9870	0.9988	0.983573	0.8839
9	0.998641	0.999975 (*)	0.9956	0.9995	0.995216	0.8813
10	0.99795	0.999563 (*)	0.9995	0.9617	0.99378	0.8102
11	0.998343	0.999213 (*)	0.9951	0.9994	0.995837	0.8782
12	0.997536	0.999313 (*)	0.9875	0.9993	0.993881	0.8577
13	0.997746	0.9997 (*)	0.9963	0.9990	0.994047	0.8875
14	0.99743	0.977017	0.9629	0.9897	0.997382 (*)	0.8079
15	0.998642	0.999698 (*)	0.9943	0.9982	0.995894	0.8941
16	0.998205	0.999759 (*)	0.9993	0.9998	0.994452	0.8677
17	0.990863	0.993312 (*)	0.9924	0.9917	0.987772	0.8581
18	0.997776	0.999914 (*)	0.9992	0.9999	0.993923	0.8684
19	0.998171	0.999307 (*)	0.9963	0.9969	0.995429	0.9010
20	0.996645	0.99936 (*)	0.9992	0.9987	0.99092	0.8472

Table 5b
Standard deviations of residuals

Data Sets	Gompertz model	Power model: general	Power model: linear (γ = 1 fixed)	Power model: multiplicative (α = 0 fixed)	Logistic model	Exponential model
1	1299.33	666.287 (*)	3095.14	1989.849267	2162.55	16019.194251
2	4243.38 (*)	5656.4	22465.7	11060.27108	4858.78	77978.042288
3	2310.44 (*)	2799.88	7900.65	52129.473267	2588.42	92281.424212
4	8435.31 (*)	11912.6	15359.7	46803.784435	15830.0	140640.4486
5	487.792 (*)	1711.06	1790.04	3760.180072	1158.13	16913.935028
6	808.405 (*)	989.635	3871.1	1091.209421	1191.26	17350.576171
7	3142.98	3016.19 (*)	6540.97	5490.518658	3881.63	22929.08106
8	103878.0	79289.1 (*)	203305.0	146408.635820	196274.0	797428.27052
9	88913.2	12028.0 (*)	165758.0	69746.738348	166088.0	1271776.1191
10	5104.65	2365.69 (*)	2598.33	12919.576864	8853.22	72810.14272
11	12638.6	8732.2 (*)	22313.7	13604.795653	19966.9	156728.413080
12	47208.0	25055.7 (*)	109793.0	25532.121358	74168.1	564910.57615
13	71584.9	26182.6 (*)	94504.8	83000.123940	116465.0	695996.77417
14	7463.57	22329.9	29148.1	35325.878563	7273.6 (*)	106499.08954
15	7388.76	3497.39 (*)	15625.7	14167.620177	12808	92138.027787
16	14011.5	5151.56 (*)	9299.4	6101.566973	24439.6	175511.480780
17	4194.58	3590.53 (*)	3930.9	4658.282097	4848.68	16243.570622
18	1996.51	394.84 (*)	1252.15	498.254165	3292.14	21096.243051
19	400.505	246.977 (*)	586.249	779.675097	631.482	3981.453854
20	1963.6	859.49 (*)	957.0	2192.549991	3218.08	19416.202077

4. The Gompertz model seems to be the right growing model for the social sciences and humanities, while the power model ($\gamma > 1$: convex growth) seems to be best for science and technology (at least as far as online databases are concerned).

We also note that our calculations are in contradiction to the ones of *Wolfram, Chu* and *Lu*.[8] Even the simple linear regression lines contradict.[8] In their study R^2-suggests that the linear model fits much better than the power model for 6 data sets when the parameter a is free; when a is fixed, they observed that the linear model fits best for 8 data sets. From a purely logical point of view, this cannot be the case since the linear model is included in the power model.

On the other hand, we have observed that the power model fits much better than all other models in 14 data sets mostly in science and technology and Gompertz model fits much better than all other models in 5 data sets, mostly in social sciences.

We will now see if the above conclusions are in agreement with our classification.

IV. Application of the classification to the 20 data sets

We calculated the 19 values of $\alpha_1(t)$ ($t = 1,....,19$) and the 10 values of $\alpha_2(t)$ ($t = 1,....,10$). The results are in Tables 6 and 7.

The conclusions (also in comparison with the statistical results) are given in Table 8. Note that we do not mention α_1 in this table: all 20 tables of α_1 are decreasing and hence of type III: *so, in our study only α_2 decides the growth model C!* If this is so in all other growth studies, then we can conclude that (based on Table 1):

– only α_2 must be calculated;

– only logistic, Gompertz, power or Ware's model are possible growth models.

Our study reveals even that Ware's model is not suitable to describe the growth of databases.

Table 6
The values of $\alpha_1(t)$

t	Set 1	Set 2	Set 3	Set 4	Set 5
1	2.147135	1.795381	1490.666667	2.185926	2.541265
2	1.591343	2.637751	2.115994	2.203409	1.932096
3	1.416188	1.546156	1.519091	1.759841	1.563196
4	1.261876	1.414845	1.391843	1.447693	1.393102
5	1.241221	1.298600	1.295110	1.309068	1.285986
6	1.209634	1.281916	1.227865	1.227071	1.270559
7	1.177132	1.492308	1.187048	1.196876	1.219080
8	1.147106	1.343190	1.167905	1.159661	1.224800
9	1.120907	1.257160	1.158344	1.133220	1.187978
10	1.118747	1.197565	1.147997	1.118850	1.160551
11	1.123626	1.167057	1.198883	1.079007	1.129906
12	1.110689	1.169626	1.177398	1.068857	1.104843
13	1.107019	1.148526	1.163569	1.059219	1.070567
14	1.102438	1.150487	1.085744	1.056695	1.062259
15	1.088657	1.148721	1.133052	1.054708	1.058169
16	1.082175	1.143527	1.073403	1.054016	1.056082
17	1.074895	1.120848	1.073571	1.049175	1.052522
18	1.075483	1.076549	1.067807	1.045424	1.048473
19	1.062527	1.020873	1.058642	1.039961	1.0408533

t	Set 6	Set 7	Set 8	Set 9	Set 10
1	4.619565	2.000651	2.287881	2.131176	3.318554
2	1.747059	1.581191	1.583432	1.575393	1.778114
3	1.433670	1.352079	1.407362	1.397740	1.382637
4	1.350517	1.254157	1.301465	1.299821	1.274901
5	1.258934	1.174616	1.228037	1.245145	1.234769
6	1.186132	1.149046	1.174949	1.205236	1.205760
7	1.214336	1.146663	1.162319	1.170933	1.160626
8	1.181731	1.133668	1.143158	1.153663	1.133778
9	1.169243	1.163338	1.138502	1.140523	1.122831
10	1.156371	1.179709	1.120602	1.124874	1.122480
11	1.161695	1.136670	1.104750	1.112905	1.104753
12	1.141281	1.112588	1.106447	1.104152	1.094118
13	1.127911	1.093243	1.110130	1.095646	1.082932
14	1.119667	1.087232	1.099106	1.087697	1.077248
15	1.107289	1.074491	1.071584	1.081532	1.070612
16	1.095486	1.074031	1.090119	1.076416	1.065866
17	1.073687	1.074456	1.096482	1.071408	1.060487
18	1.076130	1.090897	1.093959	1.069150	1.049536
19	1.049004	1.057678	1.091648	1.061462	1.038199

Table 6 (cont.)

t	Set 11	Set 12	Set 13	Set 14	Set 15
1	2.176804	2.855010	2.034805	1.959688	2.052625
2	1.623324	1.745961	1.511164	1.602327	1.525433
3	1.367986	1.496158	1.347697	1.385258	1.349709
4	1.272833	1.363219	1.263127	1.321448	1.248696
5	1.228156	1.273213	1.212826	1.227057	1.213291
6	1.187735	1.210015	1.179888	1.197249	1.182713
7	1.156161	1.187075	1.161850	1.159973	1.160721
8	1.130979	1.158747	1.140792	1.154593	1.144239
9	1.118000	1.142950	1.124365	1.142337	1.133453
10	1.127416	1.137351	1.113198	1.104966	1.119738
11	1.118378	1.120698	1.104845	1.070204	1.112140
12	1.108317	1.114432	1.094732	1.068115	1.104158
13	1.105107	1.106924	1.087305	1.085462	1.097741
14	1.096345	1.091074	1.083282	1.051313	1.090632
15	1.079985	1.101705	1.080248	1.028860	1.080806
16	1.079170	1.096987	1.076811	1.011684	1.081829
17	1.067295	1.087775	1.073542	1.006590	1.074684
18	1.056956	1.082155	1.070926	1.003214	1.068196
19	1.043586	1.081095	1.067595	1.003343	1.052916

t	Set 16	Set 17	Set 18	Set 19	Set 20
1	2.034653	2.277362	2.125442	1.926376	2.088905
2	1.556715	1.482414	1.549293	1.454304	1.614149
3	1.383410	1.313766	1.375738	1.248400	1.371643
4	1.306565	1.213659	1.257049	1.261212	1.296105
5	1.238675	1.128218	1.206379	1.206821	1.242897
6	1.181285	1.072694	1.164386	1.178112	1.202922
7	1.146492	1.121544	1.159047	1.140057	1.179710
8	1.126009	1.080885	1.133563	1.140043	1.148584
9	1.121997	1.070854	1.116380	1.127474	1.106203
10	1.114364	1.092295	1.109142	1.110206	1.084691
11	1.101701	1.094849	1.099176	1.101337	1.085001
12	1.097821	1.076254	1.090190	1.083979	1.079514
13	1.092425	1.100552	1.088428	1.085304	1.079004
14	1.078443	1.089165	1.078808	1.083588	1.069487
15	1.070885	1.088375	1.076071	1.097296	1.065330
16	1.065189	1.067340	1.065968	1.080319	1.062900
17	1.059267	1.075734	1.063433	1.068089	1.058767
18	1.056245	1.057803	1.061477	1.060289	1.062450
19	1.046594	1.036007	1.052059	1.033957	1.048549

Table 7
The values of $\alpha_2(t)$

t	Set 1	Set 2	Set 3	Set 4	Set 5
1	1.591373	2.637751	2.125894	2.203409	1.932096
2	1.787054	2.187571	2.114337	2.547709	2.177690
3	1.894611	2.455288	2.213336	2.325456	2.276218
4	2.027365	3.336802	2.204617	2.229524	2.439655
5	2.048256	3.868543	2.263627	2.159414	2.615559
6	2.113218	4.119332	2.602284	2.029598	2.569877
7	2.190931	3.647461	2.769529	1.8998001	2.397313
8	2.250163	3.567092	2.884102	1.819469	2.187320
9	2.320673	3.423752	2.854282	1.761047	2.031854

t	Set 6	Set 7	Set 8	Set 9	Set 10
1	1.747059	1.581191	1.583432	1.575393	1.778114
2	1.936195	1.695719	1.831632	1.816811	1.762725
3	2.016674	1.692720	1.877859	1.950634	1.898117
4	2.142857	1.754506	1.917176	2.027231	1.959148
5	2.301402	2.049932	1.991755	2.088773	1.999743
6	2.572435	2.256165	2.072104	2.129638	2.004670
7	2.675282	2.338699	2.175204	2.167463	2.014963
8	2.746115	2.380718	2.222767	2.187220	2.028027
9	2.71366	2.398692	2.241869	2.196757	2.010306

t	Set 11	Set 12	Set 13	Set 14	Set 15
1	1.623324	1.745961	1.511169	1.602327	1.525433
2	1.741217	2.039591	1.702312	1.830546	1.685377
3	1.856712	2.100186	1.807534	1.941329	1.791848
4	1.907419	2.119134	1.896689	1.967550	1.905852
5	1.957574	2.163612	1.957388	2.025803	1.993627
6	2.042920	2.233216	2.006529	1.934181	2.069928
7	2.140840	2.272088	2.034174	1.902813	2.135038
8	2.206161	2.369754	2.074172	1.715406	2.181701
9	2.226059	2.440643	2.120880	1.516418	2.209650

Table 7 (cont.)

t	Set 16	Set 17	Set 18	Set 19	Set 20
1	1.556715	1.482414	1.549293	1.454304	1.614149
2	1.807515	1.594464	1.729370	1.574497	1.777793
3	1.911804	1.468810	1.765766	1.793154	1.937817
4	1.888970	1.467117	1.845557	1.847896	2.025867
5	1.906717	1.521045	1.894275	1.916657	1.955763
6	1.952212	1.670842	1.949466	1.942221	1.904308
7	2.006064	1.785760	1.974960	2.003492	1.862775
8	2.032228	1.919222	1.998470	2.083255	1.836427
9	2.026519	2.039411	2.020720	2.092511	1.867445

Note on Table 7:
While computing $\alpha_2(t)$, in order to have a uniformity in all the cases, we started from $t = 1$ since we have observed irregularities in C(0): for instance in set 3, set 6, etc.. However, while fitting the various models, C(0) has been included.

We can conclude (see Table 8) that in 17 out of 20 data sets, our conclusion agrees with the statistical technique: only set 3, 6 and 20 give a different conclusion, but in all those cases the model given by table 1 is also very close in the statistical sense. Besides, it is our conviction that also in these cases the model chosen via table 1 is the more natural one, since it yields the same growth rates as the data. Finally we note the large irregularity of C(0) (observed) in data set 6 and especially in data set 3, in which we doubt that C(0) represents the actual number of publications in this year, hence exhibiting a nonnatural growth process.

V. Difference aspects of the various growth models

V. 1. First difference study

A problem that is left in this paper is the qualitative difference between the logistic model and Gompertz model for $0 < A, B < 1$. This problem was also recognised in Ref. 2 where the following result is presented.

Table 8
Qualitative conclusions versus statistical conclusions

Set	α_2	Conclusion (based on Table 1)	Based on statistical analysis (R^2 or res.st.dev.)
1	↑	Power	Power
2	↑↓	Logistic or Gompertz	Gompertz 2nd Logistic
3	↑	Power	Gompertz (close to Power)
4	↑↓	Gompertz or Logistic	Gompertz
5	↑↓	Gompertz or Logistic	Gompertz 2nd Logistic
6	↑	Power	Gompertz (close to Power)
7	↑	Power	Power
8	↑	Power	Power
9	↑	Power	Power
10	↑ (except 1 value)	Power	Power
11	↑	Power	Power
12	↑	Power	Power
13	↑	Power	Power
14	↑↓	Gompertz or Logistic	Logistic (Gompertz 2nd)
15	↑	Power	Power
16	↑ (except 2 values)	Power	Power
17	↑	Power	Power
18	↑	Power	Power
19	↑	Power	Power
20	↑↓	Logistic or Gompertz	Power (Gompertz 2nd)

Definition:

Let f: $[0, \infty[\rightarrow [0, \infty[$ be any function. Then the 'first difference function' of f is the function

$$g(t) = f(t+1) - f(t)$$

Proposition[2]*:*

(i) Let C be Gompertz model. Then the ratio (of time t versus time t+1) of the first differences of log C is a constant.
In a formula:

$$(\log C(t+2)-\log C(t+1))/(\log C(t+1)-\log C(t)) = \text{constant}$$

(ii) Let C be the logistic model. Then the ratio (of time t versus time t+1) of the first differences of 1/C is a constant:

$$\{[1/C(t+2)]-[1/C(t+1)]\}/\{[1/C(t+1)]-[1/C(t)]\} = \text{constant}$$

Proof:

(i) We have, for $C(t) = DA^{B^t}$:

$$[\log C(t+2)-\log C(t+1)]/[\log C(t+1)-\log C(t)] =$$
$$[(B^{t+2}-B^{t+1})\log A]/[(B^{t+1}-B^t)\log A].$$

Hence, since A, B $\neq$ 1, we find

$$\{[\log C(t+2)]-[\log C(t+1)]\}/\{[\log C(t+1)]-[\log C(t)]\} = B \qquad (27)$$

(ii) We have, for $C(t) = 1/(k+ab^t)$

$$[1/C(t+2)-1/C(t+1)]/[1/C(t+1)-1/C(t)] = [ab^{t+2}-ab^{t+1}]/[ab^{t+1}-ab^t]$$

Hence, since a $\neq$ 0:

$$[1/C(t+2)-1/C(t+1)]/[1/C(t+1)-1/C(t)] = b \qquad (28)$$

□

To this proposition, we can add the even simpler:

Proposition:

Let C be the Ware function

$$C(t) = \delta(1-\varphi^{-t}).$$

Then the ratio (of time t versus time t+1) of the first differences of C is a constant.

Proof:

Indeed:

$$[C(t+2)-C(t+1)]/[C(t+1)-C(t)] =$$
$$[\delta(1-\varphi^{-t-2})-\delta(1-\varphi^{-t-1})]/[\delta(1-\varphi^{-t-1})-\delta(1-\varphi^{-t})]$$

Hence, since $\delta \neq 0$,

$$[C(t+2)-C(t+1)]/[C(t+1)-C(t)] = \varphi. \square \tag{29}$$

V. 2. Differential equations

One way of 'explaining' models is to describe the differential equation where these models come from. We start with the obviously known case of the exponential distribution.

V. 2. 1. Exponential model

The growth model C is the exponential distribution if there is a constant $a > 0$ such that

$$dC(t)/dt = a\,C(t) \tag{30}$$

In this case the model is:

$$C(t) = C(0)\,e^{at} = c\,g^{t}$$

(cf. (3) or (12)). This means that the change in C ('the growth') is proportional to C itself.

V. 2. 2. Logistic model

The growth model C is the logistic model if there is a constant $a > 0$ and $K > 0$ such that

$$dC(t)/dt = a(1-C(t)/K)C(t) \tag{31}$$

In this case, the model is

$$C(t) = K/(1+(K/C(0)-1)e^{-at}) = 1/(k+ab^t)$$

(cf. (4) or (15)). This means that the proportionality factor a (of the exponential model) is now diminishing over time:
$a(1-C(t)/K)$, giving rise to the S-shaped curve.

V. 2. 3. Gompertz model

The differential equation for the Gompertz model is now

$$dC(t)/dt = (\log B)\, C(t)\, (\log C(t)-\log D) \tag{32}$$

giving:

$$C(t) = D.A^{B^t}$$

V. 2. 4. Ware's model

The differential equation for Ware's function is

$$dC(t)/dt = (\delta-C(t)) \log \varphi \tag{33}$$

giving

$$C(t) = \delta(1-\varphi^{-t}).$$

Finally

V. 2. 5. Power model

The differential equation for the power model is

$$[dC(t)/dt] - [(\gamma/t)C(t)] + \alpha\gamma/t = 0 \tag{34}$$

giving

$$C(t) = \alpha + \beta t^{\gamma}.$$

VI. Conclusions

We gave a detailed growth rate classification of the growth models C: exponential, logistic, power, Gompertz and Ware and showed that the functions α_1 and α_2 suffice to characterise C. Only the rough properties of α_1 and α_2 ($\uparrow$ or $\downarrow$) are needed to do so, so that no calculations are needed.

On the other hand, the above classification, together with nonlinear regression methods yields a sound basis to determine the exact growth model.

We illustrated our methods to the 20 online data sets of *Wolfram*, *Chu* and *Lu*[8] and proved that:

- The power law ($\gamma > 1$, convex growth) is best for modelling the growth of Sci Tech databases;
- the Gompertz function ($\log A \log B > 0$, $0 < A, B < 1$) is best for modelling the growth of social sciences and humanities databases;
- the exponential, logistic and Ware models do not fit very well;
- most of the calculations in Ref. 8 are quite different from that what we have observed, even for the simple linear regression analysis!

References

1. B.C. BROOKES, The growth, utility and obsolescence of scientific periodical literature. Journal of Documentation, 26(4), (1970) 283-294.

2. F.E. CROXTON, D.J. COWDEN, *Applied General Statistics*, Pitman & Sons, London, 1955.
3. D.J. DE SOLLA PRICE, *Little Science, Big Science*, Columbia University Press, New York and London, 1963.
4. L. EGGHE, I.K. RAVICHANDRA RAO, Citation age data and the obsolescence function: Fits and explanations *Information Processing and Management*, 28 (2) (1992) 201-217.
5. L. EGGHE, R. ROUSSEAU, *Introduction to Informetrics*, Elsevier, Amsterdam, 1990.
6. J. TAGUE, J. BEHESHTI, L. REES-POTTER, The law of exponential growth: evidence, implications and forecasts, *Library Trends*, 30(1) (1981) 125-149.
7. G.O. WARE, A general statistical model for estimating future demand levels of data-base utilization within an information retrieval organisation, *Journal of the American Society for Information Sciences*, 24 (1973) 261-164.
8. D. WOLFRAM, C.M. CHU, X. LU, Growth of knowledge: bibliometric analysis using online database data. *Informetrics 89/90*, L. EGGHE, R. ROUSSEAU (Eds) 1990. Proceedings of the second international Conference on Bibliometrics, Scientometrics and Informetrics, London (Canada), 1989, 355-372.