Available online at www.sciencedirect.com

SCIENCE DIRECT®

International Journal of Food Microbiology 96 (2004) 289–300

INTERNATIONAL JOURNAL OF Food Microbiology

www.elsevier.com/locate/ijfoodmicro

Statistical evaluation of mathematical models for microbial growth

S. López[a,*], M. Prieto[b], J. Dijkstra[c], M.S. Dhanoa[d], J. France[e]

[a] *Department of Animal Production, University of León, E-24071 León, Spain*
[b] *Department of Food Hygiene and Technology, University of León, E-24071 León, Spain*
[c] *Wageningen Institute of Animal Sciences (WIAS), Animal Nutrition Group, Wageningen Agricultural University, Marijkeweg 40, 6709 PG Wageningen, The Netherlands*
[d] *Institute of Grassland and Environmental Research, Plas Gogerddan, Aberystwyth, Dyfed SY23 3EB, UK*
[e] *Department of Animal and Poultry Science, University of Guelph, Guelph, Ontario, Canada N1G 2W1*

Received 24 October 2003; received in revised form 20 February 2004; accepted 30 March 2004

Abstract

The aim of this study was to evaluate the suitability of several mathematical functions for describing microbial growth curves. The nonlinear functions used were: three-phase linear, logistic, Gompertz, Von Bertalanffy, Richards, Morgan, Weibull, France and Baranyi. Two data sets were used, one comprising 21 growth curves of different bacterial and fungal species in which growth was expressed as optical density units, and one comprising 34 curves of colony forming units counted on plates of *Yersinia enterocolitica* grown under different conditions of pH, temperature and CO_2 (time-constant conditions for each culture). For both sets, curves were selected to provide a wide variety of shapes with different growth rates and lag times. Statistical criteria used to evaluate model performance were analysis of residuals (residual distribution, bias factor and serial correlation) and goodness-of-fit (residual mean square, accuracy factor, extra residual variance *F*-test, and Akaike's information criterion). The models showing the best overall performance were the Baranyi, three-phase linear, Richards and Weibull models. The goodness-of-fit attained with other models can be considered acceptable, but not as good as that reached with the best four models. Overall, the Baranyi model showed the best behaviour for the growth curves studied according to a variety of criteria. The Richards model was the best-fitting optical density data, whereas the three-phase linear showed some limitations when fitting these curves, despite its consistent performance when fitting plate counts. Our results indicate that the common use of the Gompertz model to describe microbial growth should be reconsidered critically, as the Baranyi, three-phase linear, Richards and Weibull models showed a significantly superior ability to fit experimental data than the extensively used Gompertz.

Keywords: Microbial growth; Growth curves; Mathematical models; Nonlinear equations; Sigmoidal functions

1. Introduction

Mathematical modelling of microbial growth has been used to estimate parameters (specific growth rate and lag time) required to study growth under different

* Corresponding author. Tel.: +34-987-291-291; fax: +34-987-291-311.
E-mail address: dp1slp@unileon.es (S. López).

0168-1605/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijfoodmicro.2004.03.026

physical and chemical conditions, to enable the effects of antimicrobials to be investigated, to formulate appropriate microbiological media or to build up prediction models for use in food and fermentation microbiology (McMeekin et al., 1993; Whiting and Buchanan, 1997). This modelling can be applied at various levels. A primary level model is an equation or function that is used to describe the microbial response over time with a characteristic set of parameter values (Whiting, 1995; McMeekin and Ross, 2002). Microbial response has been mostly expressed in terms of microbial numbers (concentration of colony forming units) or optical density as an indirect measurement (McMeekin et al., 1993).

Mathematical models can be either empirical or mechanistic (France and Thornley, 1984; McMeekin and Ross, 2002). Mechanistic models are preferred because they are derived to represent the biochemical processes controlling microbial growth. However, if the mechanism governing the process is unknown, mathematical functions have to be used empirically, and model suitability is evaluated from its ability to fit experimental data based on statistical criteria. Empirical models can develop into more mechanistic models as more information concerning the system becomes available (Baranyi and Roberts, 1995).

There are a number of sigmoidal functions that have been used for modelling somatic growth and population dynamics (France and Thornley, 1984), which could be applied to microbial growth. Despite the number of different nonlinear equations used as growth functions, there is not one growth function that is essentially superior to all others. Although there has been considerable effort to derive alternative models for microbial growth, and there are some comparative studies reported in the literature (Zwietering et al., 1990; Buchanan et al., 1997; Schepers et al., 2000; Dalgaard and Koutsoumanis, 2001), it is timely to investigate their statistical ability to fit experimental data.

The objective of the present study was to compare different models by fitting them to curves recorded under different experimental conditions (maintained constant over time), representative of an array of microbial species, growth conditions and curve shapes, and with a number and distribution of incubation points that are currently used by most researchers.

2. Material and methods

2.1. Mathematical considerations

Exponential microbial growth is represented by the general equation:

$$dN/dt = \mu N, \tag{1}$$

where N denotes microbial biomass (in units of mass, optical density, numbers, etc.), t is time (h) and μ is the specific growth rate (h^{-1}), which can be a function of time. The rate of change in the log microbial biomass [ln(N)] is then:

$$\frac{d(\ln N)}{dt} = N^{-1}\frac{dN}{dt} = \mu(t). \tag{2}$$

If L_0 denotes the value of ln(N) at time zero, then integrating Eq. (2) yields

$$\ln(N) = L_0 + \int_0^t \mu(t)dt = f(t), \tag{3}$$

giving ln(N) as a function of time.

The maximum value of μ, μ_{max}, occurs at time t^* (h), which is found by solving:

$$d\mu/dt = 0. \tag{4}$$

The lag time, T (h), is the intercept of the tangent to the steepest part of the $f(t)$ vs. time curve with the curve's lower bound. The equations of the tangent and the lower bound, respectively, are:

$$y - f(t^*) = \mu_{max}(t - t^*), \tag{6}$$

and

$$y = L_0, \tag{7}$$

giving:

$$T = t^* - \frac{f(t^*) - L_0}{\mu_{max}}. \tag{8}$$

A number of candidates for $f(t)$, presented in Table 1, will be investigated as microbial growth functions. In this table, λ (h^{-1}) and d ($h^{-0.5}$) are rate parameters,

Table 1
Candidate functions

	$f(t)$	
Linear LIN	$L_0,$	$t \leq \tau$
	$L_0 + \mu(t-\tau),$	$\tau < t < t_f$
	$L_\infty,$	$t \geq t_f$
Logistic LOG	$\frac{L_\infty}{1+e^{-\lambda(t-t^*)}}$	
Gompertz GMP	$L_\infty \exp[-e^{-\lambda(t-t^*)}]$	
von Bertalanffy VNB	$L_0 + (L_\infty - L_0)(1-e^{-\lambda t})^{1/\upsilon}$	
Richards RCH	$\frac{L_\infty}{[1+\upsilon e^{-\lambda(t-t^*)}]^{1/\upsilon}}$	
Morgan MRG	$\frac{L_0 K^\upsilon + L_\infty t^\upsilon}{K^\upsilon + t^\upsilon}$	
Weibull WBL	$L_\infty - (L_\infty - L_0)\exp[-(\lambda t)^\upsilon]$	
France FRN	$L_0,$	$t < \tau$
	$L_\infty - (L_\infty - L_0)e^{[-\lambda(t-\tau)+d(\sqrt{t}-\sqrt{\tau})]},$	$t \geq \tau$
Baranyi and Roberts BAR	$L_0 + \mu_{max} t + L_1 - L_2,$	
	$L_1 = \ln[e^{-\mu_{max} t} - e^{-\mu_{max}(t+T)} + e^{-\mu_{max} T}],$	
	$L_2 = \ln\left[1 + \frac{e^{\mu_{max}(t-T)} + e^{-\mu_{max} T}}{e^{(L_\infty - L_0)}}\right]$	

τ (h) represents a discrete lag time, K (h) is the time when half maximal growth is achieved, υ (dimensionless) is a shape or curvature parameter, L_∞ is the maximum log microbial population size.

2.2. Model evaluation

2.2.1. Data sets

Growth data recorded on 55 curves were used for model evaluation. Data were grouped in two sets of different characteristics. Data set 1 comprised 21 curves for different bacterial and fungal species grown under different conditions (Fig. 1). In this set, the growth curves were defined as the logarithm of the relative population size in terms of optical density units against time. Data are part of the software supplied by Transgalactic, Finland to the Bioscreen Reader, manufactured by LabSystems, Finland. Data for the ensuing bacterial species were used: *Actinomyces* spp., *Bacillus* spp., *Bifidobacterium* spp., *Lactobacillus plantarum*, *Clostridium* spp., *Enterococcus* spp., *Escherichia coli* (2 curves), *Lactobacillus* spp., *Listeria monocytogenes*, *Propionibacterium* spp., *Pseudomonas* spp., *Salmonella typhimurium*, *Serratia* spp., *Staphylococcus* spp. and *Streptococcus* spp., all grown at 37 °C. In addition, curves for the fungal species *Aspergillus* spp., *Candida* spp. and *Saccharomyces cerevisiae* grown at 6 or 37 °C were used. The number of data points was variable ranging between 19 and 102, with different frequency of sampling depending on the growth rate in each culture.

Data set 2 consisted of 34 curves (Fig. 2) of *Yersinia enterocolitica* JBL 1307 (strain obtained from J.B. Luchansky's collection, formerly at Food Research Institute, Madison, WI, USA) grown under different environmental conditions. Stock cultures were maintained on Trypticase Soy Agar slants at 4 °C. pH was adjusted with 5 M HCl to pH values 5, 5.5, 6, 6.5 and 7 ($\pm$ 0.01). Besides aerobic atmosphere, three combinations of carbon dioxide and oxygen (5/95, 15/85, and 25/75% CO_2/O_2) were used. Flasks were flushed for ten minutes with gas mixtures sterilized through a 0.2-μm membrane (Millipore). Inoculation was done in screw-capped flasks with rubber septum containing 250 ml of Brain Heart Infusion (BHI, Oxoid) with 0.1 ml of diluted microbial suspension (from cultures grown for 18 h at 30 °C in BHI) to give a concentration of approximately 1000 viable cells per millilitre. Flasks were placed in an incubator at 5, 10, 20 and 30 °C until stationary phase was reached. Samples taken aseptically from flasks were diluted in peptone water (1% peptone, 0.5% NaCl), plated in duplicate in BHI agar and incubated at 30 °C for 48 h. Each culture was maintained at constant conditions over the experimental period (no time-varying conditions). The number of data points was between 10 and 20, with variable frequency of sampling.

It is important to notice that the curves were not selected to study the growth characteristics of the different species (Data set 1) or the growth conditions of *Y. enterocolitica* (Data set 2), but to have a wide range of curves with marked shape differences (Figs. 1 and 2) to study how this high diversity of curves were fit by each mathematical model.

2.2.2. Model fitting

All the models were fit to the data by nonlinear regression using the NLIN procedure of the SAS package (SAS, 1988). Several possible starting values

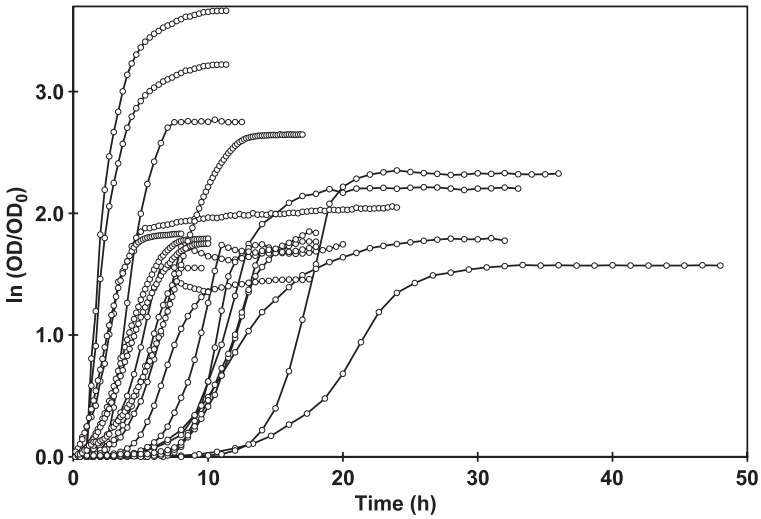

Fig. 1. Data set 1 (observed values) used for the study: Plot of 21 optical density (OD) curves of different microbial (bacteria and fungi) species [natural log of OD versus time].

were specified for each parameter, so that the NLIN procedure evaluated the model at each combination of initial values on the grid, using for the first iteration of the fitting process the combination yielding the smallest residual sum of squares (SAS, 1988). The initial values supplied were different for each data set, and the selection of the starting values was based on visual inspection of the plots. The uniqueness of the final solution achieved in each case was checked by changing the initial parameter estimates within a reasonable range for each data set.

2.2.3. Statistical analyses

To evaluate the ability of each model to describe the data without systematically over- or under-estimating any section of the curve, the number of runs of sign of the residuals (Motulsky and Ransnas, 1987) was calculated. A run is a sequence of residuals with

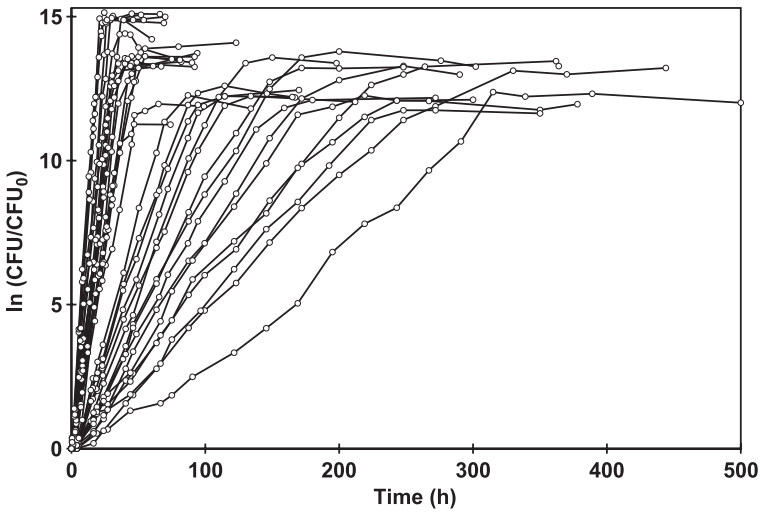

Fig. 2. Data set 2 (observed values) used for the study: Plot of 34 viable count curves of *Yersinia enterocolitica* grown under different pH, temperature and gas atmosphere conditions [natural log of number of colony forming units (CFU) versus time].

the same sign (positive or negative). As the number of observations (n) was different for each curve, the number of runs of sign was expressed as percentage of the maximum number possible (i.e. $n-1$). The probability for the occurrence of too few (indicating clustering of residuals with the same sign or systematic bias) or too many (indicating negative serial correlation) runs of sign was determined using the test for runs described by Draper and Smith (1981). The "bias factor" developed by Ross (1996) as an index of the model performance in terms of the average deviation between predicted (P) and observed (O) values was calculated as $\mathrm{BF} = \exp[\Sigma \ln(P/O)/n]$. Serial correlation was examined using the Durbin–Watson statistic, and its level of significance determined as described by Draper and Smith (1981). A number of statistics were used to evaluate the general goodness-of-fit of each model. Proportion of variation accounted for (R^2) was calculated as $1 - \mathrm{RMS}/s_y^2$, where RMS is the residual mean square and s_y^2 is the total variance of the y-variable. Average accuracy of the estimates was assessed using the "accuracy factor" of Ross (1996), as refined by Baranyi et al. (1999): $\mathrm{AF} = \exp[\sqrt{\Sigma(\ln P - \ln O)^2/n}]$. The statistical significance of the difference between models in terms of the goodness-of-fit to the same set of data was assessed by using the F-tests described by Motulsky and Ransnas (1987) for comparing two models either with the same or a different number of parameters. The difference between the residual sum of squares (RSS) of the two models and the number of parameters of each model are used to compute a F-ratio. To compare models with the same number of parameters, the following equation was used:

$$F = \frac{\mathrm{RSS}_1}{\mathrm{RSS}_2},$$

and for models with different number of parameters the equation used was:

$$F = \frac{(\mathrm{RSS}_1 - \mathrm{RSS}_2)/(df_2 - df_1)}{\mathrm{RSS}_2/df_2},$$

where df is degrees of freedom.

An alternative method for comparing models is the Akaike's Information Criterion (AIC), based on information theory (Burnham and Anderson, 2002; Motulsky and Christopoulos, 2003). For each set of data, AIC is calculated for each model as:

$$\mathrm{AIC} = n\ln\left(\frac{\mathrm{RSS}}{n}\right) + 2(p+1) + \frac{2(p+1)(p+2)}{n-p-2},$$

where n is the number of data points and p is the number of parameters of the model. The method takes into account the change in goodness-of-fit and the difference in number of parameters between two models. Comparing the individual AIC values, for each set of data, the model with the smallest AIC value is most likely to be correct (Motulsky and Christopoulos, 2003).

3. Results

The data fits obtained with all the models were evaluated statistically taking into account fitting behaviour, examination of residuals, and statistics for goodness-of-fit. In general, all the models were fit without major problems, although the initial estimates of the parameters were different for each curve in order to reach convergence within a reasonable number of iterations. If the values supplied as initial estimates were very different from the final solutions, the algorithm failed to converge in some cases. As the initial estimates were moved closer to the final solution, convergence was met in a fewer number of iterations.

Examination of the residuals obtained for each curve was based on analysis of systematic bias between observed and predicted values, clustering of residuals with the same sign and serial correlation, which may be indicative of inappropriate fitting of the model to experimental data. Perfect agreement between predictions and observations will be represented by a bias factor of 1 (Ross, 1996). Higher or lower values will indicate a systematic over- or under-estimation of the observed values, respectively. Based on the average values, the less biased models were LIN and VNB (Table 2). However, the average bias factor across the 55 curves has to be interpreted with caution as large under-estimations may balance with large over-estimations giving an average bias factor approaching unity. Therefore, the median and the

Table 2
Bias factor and Durbin–Watson statistic (DW) values obtained when fitting the models to the growth data

Model	Bias factor					No. of curves with Significant DW ($P<0.05$)	No. of curves with NON-Significant DW ($P>0.05$)
	Average	Median	Min	Max	Range		
LIN	1.008	0.985	0.919	1.330	0.410	19	32
LOG	1.104	1.057	0.962	1.497	0.535	34	15
GMP	1.043	1.001	0.905	1.339	0.433	24	19
VNB	1.010	0.998	0.884	1.284	0.400	28	17
RCH	1.060	1.023	0.965	1.329	0.364	20	25
MRG	1.022	1.002	0.920	1.295	0.375	27	17
WBL	1.023	1.005	0.939	1.293	0.354	20	24
FRN	0.955	0.951	0.506	1.385	0.880	34	15
BAR	1.018	1.020	0.944	1.092	0.148	18	30

range of values should be examined. The median of the bias factor approached unity for all the models; with model LOG showing some systematic overestimation and FRN some underestimation of the observed values. Maximum and minimum bias factors across the curves give an indication of the ability of models to fit curves with outliers. The BAR model gave the maximum bias factor closest to unity, whereas LOG and RCH gave the minimum bias factor closest to unity. It is of interest to examine the range of bias factor values (Maximum–Minimum) for each model across the curves studied, which was narrowest for the BAR model and widest for FRN. In general, bias was larger for plate count than for optical density curves.

Serial correlation of residuals was examined further with the Durbin–Watson statistic (DW) (Draper and Smith, 1981). The DW is used to test whether a model has been successful in describing the underlying trend. The DW values obtained when fitting the models to all growth data are summarized in Table 2. A DW value around 2 is statistically non-significant, and is obtained when the serial correlation is small and the residuals are distributed randomly around the zero line (when plotted against time). When the DW is significant (either its value or the difference 4-DW approaches zero), the serial correlation is significant because of the presence of cycles in the plot of residuals. With the models BAR, LIN, WBL and RCH, there were fewer curves with a significant DW and more with a non-significant DW, whereas when models LOG and FRN were used the number of curves with significant DW was greater.

The distribution of number of runs of sign is shown in Table 3. On average and across all the growth curves, the number of runs of sign was greater with models BAR and LIN, whereas with all the other models the number of runs of sign was less than 40% of the maximum for most curves. A small number of runs of sign is obtained when the residuals are not randomly distributed, so residuals of the same sign tend to cluster together on some parts of the curve. The number of cases with significantly ($P<0.05$) too few runs according to the test of Draper and Smith (1981) was similar with all the models, and only slightly smaller with model RCH (Table 3).

Table 3
Distribution of the growth curves (total number = 55) according to the computed number of runs of sign observed when fitting the curves by the different models, and number of curves that significantly showed too few number of runs of sign

Model	Number of runs of sign (as % of maximum)			No. of curves with TOO FEW Runs of sign ($P<0.05$)
	<20%	20–40%	>40%	
LIN	16	17	22	34
LOG	15	34	6	33
GMP	17	35	3	35
VNB	18	35	2	35
RCH	10	41	4	32
MRG	16	37	2	35
WBL	14	30	9	33
FRN	15	38	2	35
BAR	16	16	23	33

Table 4
Goodness of fit: R^2 values, accuracy factor and residual mean square (RMS) obtained when fitting the models to the microbial growth data, mean rank of the RMS (smallest RMS = rank 1, etc.) and number of curves (total = 55) for which the model showed the largest and the smallest RMS

	LIN	LOG	GMP	VNB	RCH	MRG	WBL	FRN	BAR
Average R^2 values	99.58	99.21	99.28	99.18	99.50	99.09	99.49	98.92	99.72
Accuracy factor									
Average	1.3070	1.5346	1.3728	1.3050	1.3293	1.3318	1.2444	1.5179	1.1616
Median	1.2253	1.3803	1.2609	1.2347	1.2128	1.2379	1.1624	1.2973	1.1579
Min	1.0351	1.0156	1.0130	1.0141	1.0078	1.0213	1.0142	1.0310	1.0229
Max	3.2646	4.2451	3.6284	3.3263	3.5419	3.5260	3.2964	3.3887	1.4249
Residual mean square									
Average	0.0583	0.2009	0.1661	0.2129	0.1547	0.2633	0.1392	0.3170	0.0560
Median	0.0386	0.1993	0.1503	0.1962	0.1435	0.2404	0.1269	0.3564	0.0339
Min	0.0014	0.0002	0.0002	0.0003	0.0001	0.0001	0.0001	0.0006	0.0001
Max	0.3798	0.7567	0.8568	1.0649	0.8177	1.2507	0.8019	1.2834	0.3541
Ranking of models according to RMS									
Mean rank RMS	3.82	5.56	5.02	6.38	3.44	6.71	3.42	8.29	2.36
No. of curves with smallest RMS	14	1	0	0	12	1	7	1	19
No. of curves with largest RMS	8	5	0	0	0	4	0	38	0

The proportion of variation explained was in general high for all the models, as the average values across the 55 growth curves were greater than 0.989 (Table 4). The R^2 values were in most cases close to unity, but some differences were detected among models, indicating that this statistic can be used as a basis for model ranking. Models BAR, LIN, WBL and RCH showed the highest average R^2 values (Table 4), whereas models MRG and FRN were the worst using this criterion. A similar trend was observed when the accuracy factor was used (Table 4), as models BAR and WBL gave values closer to unity than other models, whereas the highest average accuracy factor was observed with model FRN. The goodness-of-fit was also compared using the RMS (Table 4) that takes into account the number of parameters contained in each model. The average RMS across the 55 curves was smallest with models

Table 5
Pair-wise comparisons between models based on the number of cases where the Akaike's Information Criterion (AIC) value of the model specified in the row was smaller than that of the model specified in the column (total number of cases = 55), and ranking of models according to mean rank of the AIC (smallest AIC = rank 1, etc.) and the number of curves (total = 55) for which the model showed the largest and the smallest AIC

	LIN	LOG	GMP	VNB	RCH	MRG	WBL	FRN	BAR
Pair-wise comparisons between models									
LIN		32	40	41	33	39	32	47	16
LOG	23		24	42	17	45	19	50	13
GMP	15	31		50	32	40	34	53	6
VNB	14	13	5		1	37	5	53	4
RCH	22	38	23	54		53	21	54	20
MRG	16	10	15	18	2		7	53	6
WBL	23	36	21	50	34	48		51	16
FRN	8	5	2	2	1	2	4		4
BAR	39	42	49	51	35	49	39	51	
Ranking of models according to AIC values									
Mean rank AIC	3.91	4.76	4.25	6.60	3.82	6.69	3.93	8.49	2.55
No. of curves with smallest AIC	14	2	1	0	10	1	8	1	18
No. of curves with largest AIC	8	1	0	0	0	0	0	46	0

BAR and LIN followed by models WBL and RCH, whereas the highest average RMS was observed with models MRG and FRN. The lowest mean rank of RMS corresponded to model BAR, with model FRN ranking the highest (Table 4). The smallest RMS was observed in most curves with models BAR and LIN, whereas the greatest number of curves with the highest RMS was found with model FRN. For most plate count curves, the smallest RMS was observed with models BAR (18 out of 34) and LIN (14 out of 34). With respect to the optical density curves, the lowest RMS was found with model RCH for a significant number of curves (12 out of 21), whereas a high number of optical density curves (8 out of 21) with the largest RMS was found with model LIN. In contrast with model LIN, the fits to the BAR, WBL and RCH models never resulted in the largest RMS.

Pair-wise comparisons between models are given in Tables 5 and 6, showing that in general models BAR, LIN, WBL and RCH were superior to the other models in terms of goodness-of-fit, whereas residual variance reached with model FRN was in general significantly ($P<0.05$) larger than with other models. The other models were intermediate, and in most cases the pair-wise difference in the RMS was not large enough to reach the level of statistical significance. Therefore, when comparing models LOG, GMP, VNB and MRG, this statistical test could not detect a consistent superiority of one model over the others. The marginal column of Table 6 shows the total number of cases where each model was superior to others, showing clearly that BAR (289 cases), with LIN in second place (210 cases), has a strong edge over models such as FRN, VNB, GMP and MRG. The marginal row of this table contains the column totals, lending support to the argument that BAR, RCH and WBL have claims for consideration, as overall they were outperformed by other models for the smallest number of cases.

Table 6
Pair-wise comparisons between models: number of cases where the model specified in the row was significantly ($P<0.05$) superior (using a F-test) to the model specified in the column (total number of cases = 55)

	LIN	LOG	GMP	VNB	RCH	MRG	WBL	FRN	BAR	Total
LIN		32	39	32	20	33	17	37	0	210
LOG	14		14	15	0	9	3	22	0	77
GMP	10	6		3	0	3	2	9	4	37
VNB	9	13	1		0	0	4	5	4	36
RCH	21	35	23	19		19	15	23	11	166
MRG	11	6	8	7	1		5	14	3	55
WBL	17	26	20	14	1	12		25	10	125
FRN	5	6	2	2	1	2	4		2	24
BAR	14	45	49	44	21	44	25	47		289
Total	101	169	156	136	44	122	75	182	34	

4. Discussion

Growth models can be either empirical (set out principally to describe) or mechanistic (attempt to give a description with understanding) (France and Thornley, 1984), depending on how the model is originally derived (Baranyi and Roberts, 1994). A mechanistic model is usually derived from a differential equation relating growth rate (dN/dt) to the organism or population size (N). To be a mechanistic model, this mathematical relationship should represent the mechanism governing the growth process. This approach has been extensively used for somatic growth and population dynamics, and a large number of growth functions have been derived, such as the monomolecular, logistic, and Gompertz (Turner et al., 1976; Jason, 1983; France and Thornley, 1984).

Occasionally, a similar approach has been used for microbial growth models (Jason, 1983; McMeekin et al., 1993; Baranyi and Roberts, 1994), but with little recognition due, in part, to the form or units used to quantify microbial growth. Numbers (colony forming units) or absorbance units have been used extensively to measure microbial growth, but this sort of data require a logarithmic transformation because of their heteroscedasticity (are not normally distributed and variance is not uniform over all measurement conditions). If data are not transformed, the regression analysis will be flawed (Schaffner, 1998). Probably the most common models used for somatic growth could be applied to microbial growth if microbial mass rather than numbers or concentration was used to measure growth. However, cell numbers or optical density are used because these measurements are much simpler and faster to make, although this may deserve some attention because the cell size may change with growth conditions, resulting in cell number dynamics lagging behind biomass dynamics (Nielsen et al., 1997). Microbial numbers are also easier to

interpret in studies of infectious diseases or food microbiology.

The relationship between ln(N) and time follows a sigmoidal pattern that can be described empirically using available growth functions. Gibson et al. (1987) used Gompertz and logistic equations to fit microbial growth data with a reasonable goodness-of-fit, allowing for the estimation of descriptive growth parameters (maximum specific growth rate and lag time). However, the use of ln(N) instead of N as the response variable means that the models are not simply transformations of the original Gompertz and logistic models derived mechanistically (McMeekin et al., 1993), and thus should be called modified Gompertz and logistic models, respectively.

A differential equation for a modified Gompertz was derived by Van Impe et al. (1992), and a similar approach could be used to establish the function for $\mu(t)$ for a wide range of growth functions. However, the expressions for the specific growth rate (μ) obtained from Eq. (2) would be intricate and not really suited to mechanistic interpretation. Strictly, models used to represent growth are not wholly mechanistic, because regardless of the biological detail of their derivation, they are not derived from a set of differential equations in rate/state form constructed by direct application of scientific law. Nevertheless, as the equations seem to mimic reality, it is anticipated that the quantitative concepts and mechanisms underlying some of these expressions will be identified in the future as more understanding is gained (Box and Draper, 1987).

On the other hand, when these models have been derived for somatic growth or population dynamics, the specific growth rate (μ) is a function of somatic or population size (N), and does not depend on time directly (autonomous models) (Baranyi et al., 1993). However, in this case, and for most of the models studied, μ derived for most of the different candidate equations would be an explicit function of time, and therefore have to be considered semi- or non-autonomous differential equations to model growth (Baranyi et al., 1993).

Although the modified Gompertz model has been used most extensively in food microbiology, other models have been applied. A semi-autonomous model was derived by Baranyi and Roberts (1994) with a more rational biological interpretation. With six parameters, it can be considered a flexible, generalized model that encompasses other sigmoidal functions (Baranyi, 1997). However, sometimes this complicated model cannot be fitted by usual nonlinear regression programs due to the large number of parameters included in the model and its sensitivity to the number of data points and their distribution (Buchanan et al., 1997; Baranyi, 1997). This inconvenience detracts from more widespread application of the model. Therefore, reparameterization of the model is desirable, and a procedure has been adopted in which some assumptions are applied to simplify the original model to a more suitable four-parameter equation (Baranyi and Roberts, 1995).

Buchanan et al. (1997) resolved the debate about mechanistic derivation using the same principles as in somatic growth assuming microbial growth follows first-order kinetics, implying that growth rate at any time is proportional to the size of the microbial population at that time, with constant fractional rate μ, i.e. growth rate is a function of size or cell concentration, not of time (Jason, 1983), and thus is an autonomous model. This kinetics is represented by an exponential function for microbial size (N) and by a straight line with slope μ if the variable is log transformed. The sigmoidal shape is achieved by including a discrete lag phase at the lower asymptote and an abrupt cut off at the upper asymptote to represent the stationary phase (France and Thornley, 1984). The result is a visually simple model with three spline-lines that describes experimental data usefully (Garthright, 1997).

The present study has compared the statistical performance of a number of sigmoidal models to describe microbial growth, including the three-phase linear model of Buchanan et al. (1997), the four-parameter model derived by Baranyi and Roberts (1995), and seven other sigmoidal functions previously used to describe somatic growth (Turner et al., 1976; France and Thornley, 1984; France et al., 1996; López et al., 2000; Darmani Kuhi et al., 2002). These functions have also been used to describe rumen degradation kinetics (López et al., 1999), highlighting the close relationship between substrate degradation and microbial growth. The logistic and Gompertz are sigmoidal models with a fixed point of inflection at 1/2 and 1/e of the distance between the lower and the upper (L_∞) asymptotes,

respectively. The modified Gompertz represents an asymmetrical sigmoidal shape and may offer greater flexibility than the logistic (Gibson et al., 1987). The other models provide robust flexible growth functions, capable of describing both diminishing returns and sigmoidal behaviour, and with a variable point of inflection that can occur at any growth stage between L_0 and L_∞. The flexible functions encompass other growth equations as special or simpler cases; for example, RCH encompasses LOG, GMP and VNB (France and Thornley, 1984; Zwietering et al., 1990; Dalgaard and Koutsoumanis, 2001). Except for the FRN model, the point of inflection can be calculated using simple algebraic expressions.

The modified Gompertz model (Gibson et al., 1987) has been accepted commonly and used extensively, mainly on the grounds of the results reported by Zwietering et al. (1990), who concluded that the modified Gompertz equation was statistically sufficient to describe microbial growth data and was easy to use. With the adoption of this model by the FoodMicroModel consortium in the UK and the Pathogen Modelling Program group at USDA, the Gompertz function has been extensively used to achieve comparability and consistency in global databases. However, several studies have concluded that the modified Gompertz equation shows systematic lack of fit to microbial growth data (Baranyi et al., 1993; Dalgaard, 1995; Dalgaard et al., 1994; Membré et al., 1999; Schepers et al., 2000).

There is not a single, simple statistical method to evaluate similarities and differences between nonlinear models, and to deal with the question of which model should be used. Usually, a number of procedures are used to obtain an overall view of model behaviour and choose the one that is more consistent for most of the tests performed. The main statistical procedures used for model comparison are residual analysis and tests for goodness-of-fit (Draper and Smith, 1981; Motulsky and Ransnas, 1987; Motulsky and Christopoulos, 2003).

Statistical performance led to some suggestions for the evaluation of models, as most of the statistical tests performed were consistent in rating the models, although an overall assessment is required. According to the analyses of residuals, the BAR, LIN, WBL and RCH models showed the best behaviour. In general, the number of runs of sign was small with all the models, especially for the optical density curves, but the deviation between predicted and observed values cannot be large considering the average bias factor. The narrow range of bias factor values observed for the BAR model (Table 2) reveals its consistency, resulting in values close to unity for most curves, and may be an important argument for its use, whereas the wide range observed for FRN is evidence of poor performance. Some serial correlation can be expected with growth data, because size at time $t-1$ is very likely to be autocorrelated with size at time t when measurement intervals are short. A lower percentage of curves with significant autocorrelation was observed when BAR, LIN, WBL and RCH models were fit, followed by GMP.

Criteria for choosing a model that have proved valid are statistical goodness-of-fit and convenience. The difficulty of discriminating between models based on the statistical fitting may be because most statistical tests for goodness-of-fit have been designed for linear models and are of limited value for comparing nonlinear models (Draper and Smith, 1981; Motulsky and Ransnas, 1987). Most statistical tests revealed a general trend with some significant differences between models in their goodness-of-fit. Generally, all the models studied proved valid for most growth curves, but it can be concluded that the fit with the BAR, LIN, WBL and RCH models was significantly better than with the others, showing the highest R^2 values, best accuracy factor and lowest residual variance and AIC values. The smallest RMS and AIC values were observed with the BAR model in 33% and 35% of the curves fitted, respectively. The mean rank based on either RMS or AIC values was noticeably better with BAR than with any other model, followed by LIN, RCH and WBL. The BAR, LIN and RCH models improved significantly ($P<0.05$) the goodness-of-fit in comparison with the Gompertz for 89%, 71% and 42% of the growth curves, respectively. Based on pair-wise comparisons (Tables 5 and 6), the BAR model resulted in a better goodness-of-fit for a greater number of curves than any other model, resulting in the lowest mean rank of RMS and AIC values. The LIN, RCH and WBL models showed an acceptable overall performance, and although were not superior to the BAR model, their behaviour across all the curves was more consistent than that of the other models. The RCH model was the best for the

optical density curves, and overall was virtually never outperformed by any model other than LIN or BAR (Table 6). Buchanan et al. (1997) showed that the goodness-of-fit attained with the three-phase linear model compared well with established models (Gompertz and Baranyi). However, if only the optical density curves were considered, the LIN model resulted in the worst fit for 8 out of the 21 curves. The large number of data points in these curves may lead to a steady evolution in the lag period and a slow transition from the growth phase to the stationary phase. The performance of the LIN model is better when both changes are rapid, but this model seems to be more limited than the others when describing smoother profiles as with the optical density curves (Baranyi, 1997).

Based on the overall statistical analysis performed, it is evident that model FRN showed some deficiencies that were detected by most tests. The other four-parameter flexible functions (VNB and MRG) did not show any advantage over the LOG or GMP models. The behaviour of these latter models was similar (Giannuzzi et al., 1998), without any consistent advantage in favour of either of them. The extensively used Gompertz model may be satisfactory to fit growth curves, but in contrast with previous results (Gibson et al., 1987; Zwietering et al., 1990), other models are shown to be more accurate for fitting microbial growth curves (Baranyi et al., 1993; Dalgaard, 1995; Dalgaard et al., 1994; Membré et al., 1999; Schepers et al., 2000). Fitting the GMP model was associated with higher serial correlation and, in terms of goodness-of-fit, other models did better than this classical function, in particular the Baranyi and the Richards models, in agreement with other authors (Schepers et al., 2000; Dalgaard and Koutsoumanis, 2001). Zwietering et al. (1990) concluded that the modified Gompertz model seemed to be sufficient, compared with the flexible Richards model, to fit microbial growth curves. In the present study, a wider range of microbial curves with a high diversity of shapes and profiles using both optical density and microbial number curves were analyzed, so it might be expected that the results are more widely applicable. The large diversity of growth curves from a variety of species considered herein (see Figs. 1 and 2) would strongly suggest that the mathematical functions selected can be applied to the growth curves (microbial biomass vs. time) of most microbial strains.

Although the models have been tested using statistical procedures, biological implications can be derived from our results. First, it must be stressed that only models from which meaningful biological indicators can be determined have been selected, and mathematical expressions to calculate these indicators (e.g. maximum specific growth rate and lag time) are derived and presented. On the other hand, statistical goodness-of-fit of each model is a useful unbiased criterion to establish which model is superior in describing the biological/experimental data, and in yielding more reliable and accurate estimates of the growth indicators (Schepers et al., 2000). Model choice based on unbiased and objective criteria is essential, as precise estimates of maximum specific growth rate and lag time are required if the optimal or the non-favourable growth conditions are to be assessed.

In conclusion, probably all the sigmoidal functions evaluated can be used as primary level microbial growth models with an acceptable degree of goodness-of-fit. However, a detailed statistical evaluation across a wide range of curve shapes and profiles has revealed some significant differences among models in their performance and accuracy. Models BAR, LIN, RCH and WBL showed a consistent performance for both types of microbial growth curves, although it was noticeable that BAR and LIN models showed the best fit for plate count data and model RCH for optical density profiles. Nevertheless, based on only statistical criteria, it would be appropriate to select models BAR, LIN, RCH and WBL for general use in describing microbial growth curves.

Acknowledgements

The authors wish to acknowledge the financial support received from CICYT Project no. AGF98-0188. Support for J.F. from the Spanish Secretaría de Estado de Educación y Universidades del Ministerio de Educación, Cultura y Deporte and the European Social Fund (Ayuda para estancias de profesores, investigadores, doctores y tecnólogos extranjeros en España Ref. SAB2000-0112) is gratefully acknowledged.

References

Baranyi, J., 1997. Simple is good as long as it is enough. Food Microbiol. 14, 391–394.

Baranyi, J., Roberts, T.A., 1994. A dynamic approach to predicting bacterial growth in food. Int. J. Food Microbiol. 23, 277–294.

Baranyi, J., Roberts, T.A., 1995. Mathematics of predictive food microbiology. Int. J. Food Microbiol. 26, 199–218.

Baranyi, J., Roberts, T.A., McClure, P., 1993. A nonautonomous differential equation to model bacterial growth. Food Microbiol. 10, 43–59.

Baranyi, J., Pin, C., Ross, T., 1999. Validating and comparing predictive models. Int. J. Food Microbiol. 48, 159–166.

Box, G.E.P., Draper, N.R., 1987. Empirical Model-Building and Response Surfaces. Wiley, New York.

Buchanan, R.L., Whiting, R.C., Damert, W.C., 1997. When is simple good enough: a comparison of the Gompertz, Baranyi, and three-phase linear models for fitting bacterial growth curves. Food Microbiol. 14, 313–326.

Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, 2nd edition. Springer-Verlag, New York, NY.

Dalgaard, P., 1995. Modeling of microbial activity and prediction of shelf-life for packed fresh fish. Int. J. Food Microbiol. 26, 305–317.

Dalgaard, P., Koutsoumanis, K., 2001. Comparison of maximum specific growth rates and lag times estimated from absorbance and viable count data by different mathematical models. J. Microbiol. Methods 43, 183–196.

Dalgaard, P., Ross, T., Kamperman, L., Neumeyer, K., McMeekin, T.A., 1994. Estimation of bacterial-growth rates from turbidimetric and viable count data. Int. J. Food Microbiol. 23, 391–394.

Darmani Kuhi, D., Kebreab, E., López, S., France, J., 2002. A derivation and evaluation of the von Bertalanffy equation for describing growth in broilers over time. J. Anim. Feed Sci. 11, 109–125.

Draper, N.R., Smith, H., 1981. Applied Regression Analysis. Wiley, New York.

France, J., Thornley, J.H.M., 1984. Mathematical Models in Agriculture. Butterworths, London.

France, J., Dijkstra, J., Thornley, J.H.M., Dhanoa, M.S., 1996. A simple but flexible growth function. Growth Dev. Aging 60, 71–83.

Garthright, W.E., 1997. The three-phase linear model of bacterial growth: a response. Food Microbiol. 14, 395–397.

Giannuzzi, L., Pinotti, A., Zaritzky, N., 1998. Mathematical modelling of microbial growth in packaged refrigerated beef stored at different temperatures. Int. J. Food Microbiol. 39, 101–110.

Gibson, A.M., Bratchell, N., Roberts, T.A., 1987. The effect of sodium chloride and temperature on the rate and extent of growth of *Clostridium botulinum* type A in pasteurized pork slurry. J. Appl. Bacteriol. 62, 479–490.

Jason, A.C., 1983. A deterministic model for monophasic growth of batch cultures of bacteria. Antonie Van Leeuwenhoek 49, 513–536.

López, S., France, J., Dhanoa, M.S., Mould, F., Dijkstra, J., 1999. Comparison of mathematical models to describe disappearance curves obtained using the polyester bag technique for incubating feeds in the rumen. J. Anim. Sci. 77, 1875–1888.

López, S., France, J., Gerrits, W.J., Dhanoa, M.S., Humphries, D.J., Dijkstra, J., 2000. A generalized Michaelis–Menten equation for the analysis of growth. J. Anim. Sci. 78, 1816–1828.

McMeekin, T.A., Ross, T., 2002. Predictive microbiology: providing a knowledge-based framework for change management. Int. J. Food Microbiol. 78, 133–153.

McMeekin, T.A., Olley, J.N., Ross, T., Ratkowsky, D.A., 1993. Predictive Microbiology: Theory and Application. Wiley, New York.

Membré, J.M., Ross, T., McMeekin, T.A., 1999. Behaviour of *Listeria monocytogenes* under combined chilling processes. Lett. Appl. Microbiol. 28, 216–220.

Motulsky, H.J., Christopoulos, A., 2003. Fitting Models to Biological Data Using Linear and Nonlinear Regression. A Practical Guide to Curve Fitting. GraphPad Software, San Diego, CA.

Motulsky, H.J., Ransnas, L.A., 1987. Fitting curves to data using nonlinear regression: a practical and nonmathematical review. FASEB J. 1, 365–374.

Nielsen, L.K., Reid, S., Greenfield, P.F., 1997. Cell cycle model to describe animal cell size variation and lag between cell number and biomass dynamics. Biotechnol. Bioeng. 56, 372–379.

Ross, T., 1996. Indices for performance evaluation of predictive models in food microbiology. J. Appl. Bacteriol. 81, 501–508.

SAS, 1988. SAS/STAT® User's Guide (Release 6.03). SAS Inst., Cary, NC.

Schaffner, D.W., 1998. Predictive food microbiology Gedanken experiment: why do microbial growth data require a transformation? Food Microbiol. 15, 185–189.

Schepers, A.W., Thibault, J., Lacroix, C., 2000. Comparison of simple neural networks and nonlinear regression models for descriptive modeling of *Lactobacillus helveticus* growth in pH-controlled batch cultures. Enzyme Microb. Technol. 26, 431–445.

Turner, M.E., Bradley, E.L., Kirk, K.A., Pruitt, K.M., 1976. A theory of growth. Math. Biosci. 29, 367–373.

Van Impe, J.F., Nicolai, B.M., Martens, T., De Baerdemaeker, J., Vandewalle, J., 1992. Dynamic mathematical model to predict microbial growth and inactivation during food processing. Appl. Environ. Microbiol. 58, 2901–2909.

Whiting, R.C., 1995. Microbial modeling in foods. Crit. Rev. Food Sci. Nutr. 35, 464–494.

Whiting, R.C., Buchanan, R.L.B., 1997. Predictive modelling. In: Doyle, M.P., Beuchat, L.R., Montville, T.J. (Eds.), Food Microbiology Fundamentals and Frontiers. American Society for Microbiology Press, Washington, DC, pp. 728–739.

Zwietering, M.H., Jongenburger, I., Rombouts, F.M., van't Riet, K., 1990. Modeling of the bacterial growth curve. Appl. Environ. Microbiol. 56, 1875–1881.