b
UNIVERSITÄT
BERN

THEORETICAL ECOLOGY AND EVOLUTION

INSTITUTE OF ECOLOGY AND EVOLUTION

SUMMER PROJECT REPORT, 2022

Statistical comparison of sigmoidal kinetic models for descriptive modelling of microbial growth on linear, logarithmic, first, second order and logarithmic derivative scales.

Harsh Bardhan Gupta

supervised by
Prof. Dr. Claudia Bank & Dr. Loïc Marrec

Acknowledgements

I sincerely thank my project supervisor Prof. Dr. Claudia Bank, for giving me the opportunity to carry out my remote summer internship. I thank her for giving me this project and guiding me throughout the project. I would also like to thank my post-doc supervisor Dr. Loïc Marrec for helping me understand, proceed, and complete this project. I thank him for constantly observing my work and giving essential suggestions.

List of Abbreviations

OD Optical Density

G.O.F Goodness-of-fit

RSE Residual Standard Error

AdjRsq Adjusted R-Squared

F-stats F-Statistic

AIC Akaike's Information Criterion

BIC Bayesian information criterion

N Population size

t time

ln Natural log

d/dt first derivative

d2/dt2 second derivative

No. Number

μ growth rate

K population's carrying capacity.

Statistical comparison of kinetic models for fitting microbial growth data.

Harsh Bardhan Gupta

Contents

1 Introduction

A populations is composed of individuals of the same species present simultaneously in the same location and that can reproduce. **Population growth**, which is the process by which the population size changes over time as individuals reproduce, is an important characteristics studied extensively by ecologists.Population size can change due to birth, death, emigration and immigration. Understanding the causes and effects of population growth enables scientists to predict future changes in population sizes and growth rates with better accuracy. This is very helpful in various fields. One of the greatest pauses in the development of human life, society, country, or say the whole world in history is caused by deadly viruses and bacteria, the growth rate can help monitor and predict the future of the infection and any further outbreak. In microbiology, it helps in the study of growth to investigate the effects of antimicrobials, the formulation of suitable microbiological media, or the development of prediction models for application in food and fermentation microbiology. Numerous models have been derived to describe population growth. These address population dynamics in various ways, either discretely or, for huge populations, mainly continuously. Now the question arises about why the study of population growth is important and why new methods or models to describe it are emerging.

The basic trend in population growth can be described using the Exponential growth model when the condition is ideal, or the population of the area is in the initial period, i.e., there is no predation or intra-specific competition and an unlimited supply of food and resources. The straightforward exponential growth equation can offer a good estimate of such growth. However, generally, a stable population will reach a saturation level characteristic, which is typically called the carrying capacity, K, and forms a numerical upper bound on the growth size. Resources with a limited supply cannot exhibit exponential growth. The growth will therefore exhibit a lag phase, an exponential or log phase, and stationary phase, an asymptote. Sigmoid growth models are used to describe this kind of growth. This included the basic empirical models like Logistic (Vogels et al., 1975, Pearl and Reed, 1920), Gompertz (Gompertz, 1825) and Richards (RICHARDS, 1959 and many more emerging complex mechanistic models like Barnayi (Baranyi et al., 1993) and Huang (Huang, 2011). The kinetic models of all these growth models estimate the three most important parameters (carrying capacity, growth rate, and initial population size), which are needed to analyze growth under various conditions. All these models are very much different from one another in terms of equations and how they are derived. Despite the variety of nonlinear equations employed as growth functions, no single growth function is fundamentally better than all others. When used to fit growth curves, all of these models have some advantages and some disadvantages.

There have been a few comparative studies published in the literature that have compared and statistically evaluated the goodness of fit of several growth models for the N vs. t (i.e., population size vs. time) experimental data(López, Prieto, Dijkstra, Dhanoa, et al., 2004, Buchanan et al., 1997, Dalgaard and Koutsoumanis, 2001, Schepers et al., 2000). Because no matter how good a model theoretically describes the growth, it is much important to look into how well these models statistically fit experimental data, as real nature is complex. For example (López, Prieto, Dijkstra, Dhanoa, et al., 2004) evaluated which mathematical functions (among linear, logistic, Gompertz, Von Bertalanffy, Richards, Morgan, Weibull, France, and Baranyi) are most appropriate for describing the growth curves of microorganisms. To evaluate such, they have used linear (N vs. t) growth curves of several bacterial and fungus species in different stress conditions using several goodnesses of fit statistical criteria. Typically Optical density (OD) is used to indirectly measure population size (N) and study the growth of microorganisms (McMeekin and Ross, 2002). They showed that the widely used Gompertz model was significantly outperformed by the Baranyi, three-phase linear, Richards, and Weibull models when it came to fitting experimental data. Generally, mechanistic models will outperform since they were developed to represent microbial development's biochemical mechanisms. However, suppose the mechanism regulating the process is unclear. In that case, mathematical functions need to be applied empirically, and model adequacy is assessed from its capacity to fit experimental data based on statistical criteria (Baranyi and Roberts, 1995, France and Thornley, 1984). All these comparison studies were performed on growth data on a linear scale, i.e, $(ODvs.t)$, in which most of the models fit in a similar fashion very nicely. However, things start to behave differently when the scale is altered, such as logarithmic or derivative scales. A significant difference in model fit can be easily observed in $ln(OD/OD_0)\ vs\ t$, $d(OD)/dt\ vs\ t$, $d^2(OD)/dt^2\ vs\ t$, $d(ln(OD/OD_0))/dt\ vs\ t$ curves. This study aimed to compare several kinetic models by fitting them to curves obtained under various experimental conditions, typical of a variety of microbial species, growth conditions, and curve forms, as most of the previous comparison papers use statistical models rather than kinetic. We have tested several kinetic models because nature is complex and unpredictable; no one exactly knows what happens in a beaker containing bacteria. So, by comparing these, we wanted to find an overall best model which is closer to nature or, to say, experimental data (controlled nature). Additionally, the more important thing we wanted to compare was the results of fitting various models in different scales, Why certain models come the as better option to fit in certain scales.

We have made the statistical comparison between sigmoidal kinetic models logistic, Gompertz, Richards, Baranyi, and Huang. To perform this comparison, 128 data sets of different bacterial species grown under different stress conditions were used. All the data sets were adapted from different papers, and all growth was expressed in optical density units. Several statistical criteria were used to evaluate model goodness of fit, for example - Residual Standard error, Adjusted R square, F-statistic, and model comparison criteria like Akaike's Information criterion and Bayesian information criterion. For the model fitting, a similar method as (López, Prieto, Dijkstra, Dhanoa, et al., 2004) was applied using R-studio code with some crucial changes for fitting and evaluating models at a scale different from linear. Furthermore, for comparison and evaluation, similar to (López, Prieto, Dijkstra, Dhanoa, et al., 2004), the mean rank and number of the curve with the best and worst scores were calculated for all the statistical criteria in all the different scales.

2 Methods

2.1 Data sets

All 128 data sets were adapted from different papers (Adkar et al., 2017), (Todd and Selmecki, 2020) and (Hammer et al., 2021). The data set contains curves of mean OD vs. t for different strains of the same bacteria or different bacterial species grown under different conditions. It is important to note that the curves were not chosen to examine the growth characteristics of various species or the growth conditions but rather to have a wide range of curves with notable shape variations to examine how each mathematical model fits this high diversity of curves.

2.2 Sigmoid growth models

Modeling the population of a species that increases exponentially over time is performed by -

$$N(t) = N_0 e^{\mu t}$$

where N_0 is the initial population at time t=0, $N(t)$ is the population at time t, and μ is the growth rate. Smaller populations still unrestricted by their environment or the resources nearby are typically where this type of growth occurs. We can use the equation above to estimate population because growth is almost constant in small populations.

However, generally, the population will begin to approach its carrying capacity as it grows, which is the highest population that the environment can support. The population growth will then start to slow down. Growth will be negative until the population falls back to carrying capacity or below if the population is ever exceeded. We have to apply different sets of equations to model this kind of population growth, which take carrying capacity into account.

Logistic model P. F. Verhulst (Vogels et al., 1975) developed and termed a logistic curve to represent population growth and or the increase in the size of an organ or population. The kinetic form of this model is described as-

$$N(t) = \frac{N_0 K e^{\mu t}}{K + N0(e^{\mu t} - 1)}$$

where N_0 is the initial population at time t=0, $N(t)$ is the population at time t, μ is the growth rate, and K is the population's carrying capacity.

Gompertz model Actuary B. Gompertz proposed the Gompertz function utilizing the exponential function to explain the relationship between a rising death rate and age when he presented his law of human mortality (Gompertz, 1825). The kinetic form of this model is described as-

$$N(t) = (N_0/K)^{e^{-\mu t}} K$$

where N_0 is the initial population at time t=0, $N(t)$ is the population at time t, μ is the growth rate, and K is the population's carrying capacity.

Richards model The Richards model (RICHARDS, 1959), also known as the generalized logistic function, provides flexibility in the asymmetry by incorporating the fourth parameter, β, which dictates which asymptote is closest to the inflection point. The kinetic form of this model is described as-

$$N(t) = \frac{N_0 K}{(N_0^\beta + (K^\beta - N_0^\beta)e^{-\mu\beta t})^{-1/\beta}}$$

where N_0 is the initial population at time t=0, $N(t)$ is the population at time t, μ is the growth rate, K is the population's carrying capacity, and β is the additional parameter to adjust inflection point.

Baranyi model Baranyi (Baranyi et al., 1993, Baranyi and Roberts, 1994) established a new growth model family in which the physiological state of the cells is represented by a single variable. The value of that variable at inoculation and the post-inoculation environment decide how long the lag will last. The physiological state of the inoculum is largely consistent and independent of the subsequent growth circumstances when the subculturing technique is standardized, as it is in laboratory research that leads to models. The kinetic form of this model is described as-

$$N(t) = \frac{(-1+e^{h_0}+e^{\mu t})N_0 K}{(e^{\mu t}-1)N_0 + e^{h_0} K}$$

where N_0 is the initial population at time t=0, $N(t)$ is the population at time t, μ is the growth rate, K is the population's carrying capacity, and h_0 is the lag duration parameter.

Huang model For the purpose of describing microbial growth in isothermal settings, a new mechanistic growth model was created by Huang (Huang, 2011). The fundamental observation of bacterial development, which might comprise stationary, exponential, and lag phases, was used to derive the new mathematical model. Huang's model used nonlinear regression to predict the length of the lag phase and the exponential growth rate of a growth curve. The Baranyi and Huang models are nearly identical, but the presupposition of the influence of history and bacterial physiological conditions is not required in the Huang model. The model is simple and intuitive, with clearly defined stationary phases and lag exponential growth. The kinetic form of this model is described as-

$$N(t) = \frac{N_0 K}{N_0 + (K - N_0)(1+e^{\alpha\tau})^{\mu/\alpha}(e^{\alpha t}+e^{\alpha\tau})^{-\mu/\alpha}}$$

where N_0 is the initial population at time t=0, $N(t)$ is the population at time t, μ is the growth rate, K is the carrying capacity of the population, and α & τ are additional parameters.

2.3 Estimation of parameters and Model fitting

All the models were fit to the *OD vs. t* data by using nlsLM the function of the minpack.lm R-package. nlsLM is a modified version of nls (Nonlinear Least Squares) that uses nls.lm for fitting. The nls.lm function offers an R interface to the MINPACK library's lmder and lmdif functions, which are used to solve nonlinear least-squares problems by modifying the Levenberg-Marquardt algorithm and supporting lower and upper parameter constraints. Starting estimates in each model for N_0 and K parameter was fixed as the minimum of OD and median of OD, respectively. Starting estimates for μ and other extra parameters like β, *h0*, α, and τ are chosen from a combination of several possible starting values of each parameter so that the nls.lm evaluated each combination of initial values on the grid, which used the combination producing the smallest residual sum of squares for the fitting process' initial iteration. After that, predict function was used to predict the expected y value i.e, OD values in OD vs t curve based on the model equation. For derived data like $ln(OD/OD_0)\ vs\ t$, $d(OD)/dt\ vs\ t$, $d^2(OD)/dt^2\ vs\ t$, $d(ln(OD/OD_0))/dt\ vs\ t$ fitting was done simply by evaluating expected y-axis value from derived equations of models using optimal parameter obtained from $OD\ vs\ t$ data fitting. This method is adapted from (López, Prieto, Dijkstra, Dhanoa, et al., 2004) with slight changes as necessary.

2.4 Goodness-of-fit statistical measures

In linear regression models, the least squares parameter estimators are unbiased, normally distributed, and minimize variance. According to (Seber, 2003), these characteristics are the most desired ones an estimator can have. But, least squares estimators may not have the qualities mentioned above of a good estimator in nonlinear regression models, especially with very small sample sizes. Because the parameter estimates and related standard errors will be biased, the nonlinear behavior of the model parameters is undesirable as it would result in incorrect inferences. A variety of statistics were employed to assess each model's general goodness-of-fit (López, Prieto, Dijkstra, Dhanoa, et al., 2004).

Residual Standard Error Residual Standard Error was calculated, similar to Standard deviation with one difference of dividing SSE (Sum of Squares of Errors) with n minus 1 + the number of variables involved rather than by n-1.

$$RSE = \sqrt{(}SSE/(n-(1+k)))$$

where $SSE = \sum(y - y_{pred})^2$ is sum of squares errors, n is total number of data points and k is total number of model parameters.

Adjusted R-Squared The proportion of variation accounted for Adjusted R-Squared was calculated taking into account the number of samples and variables.

$$Adj.R^2 = 1 - (SSE/SSyy) * (n-1)/(n-(k+1))$$

where $SSE = \sum(y - y_{pred})^2$ is sum of squares of errors, $SSyy = \sum((y - mean(y))^2)$ is the total variance of the y-variable, n is total number of data points and k is total number of model parameters.

F-Statistic F-Statistic tests our model globally to see if it has at least one significant variable. It also considers the number of variables and observations used.

$$F.stat = ((SSyy - SSE)/k)/(SSE/(n-(k+1)))$$

where $SSE = \sum(y - y_{pred})^2$ is sum of squares of errors, $SSyy = \sum((y - mean(y))^2)$ is the total variance of the y-variable, n is total number of data points and k is total number of model parameters.

Akaike's Information Criterion The Akaike's Information Criterion (AIC), which is based on information theory, provides an alternative way for comparing models (*Model Selection and Multimodel Inference*, 2004); (Motulsky and Christopoulos, 2003). The AIC assesses the relative amount of information a given model loses; the lower the loss, the higher the model's quality.AIC addresses both the concern of overfitting and the concern of underfitting. For each set of data, AIC is calculated for each model as:

$$AIC = n * ln(MSE) + 2 * (k+1) + 2 * (k+1) * (k+2)/(n-k-2)$$

Where $MSE = mean((y_c - y)^2)$ is the mean of the squares of errors, n is the total number of data points, and k is the total number of model parameters.

Bayesian information criterion When choosing a model from a limited number of options, we use the Bayesian information criterion (BIC), also known as the Schwarz information criterion. Models with lower BIC are typically chosen (Schwarz, 1978). BIC is similar to AIC, based on the log-likelihood function. It is feasible to enhance the likelihood when fitting models by adding parameters but doing so runs the risk of overfitting. The BIC penalizes the complexity model more than the AIC, which means that more complex models will have a worse (bigger) score and will, in turn, be less likely to be chosen.

$$BIC = (n * ln(MSE) + (k * ln(n)))$$

Where $MSE = mean((y_c - y)^2)$ is the mean of the squares of errors, n is the total number of data points, and k is the total number of model parameters. Assuming that the model's errors or disturbances are independent, identical, and have a normal distribution, and assuming that the derivative of the log-likelihood with respect to the true variance is zero.

2.5 Mean Ranking of Models

First, all the growth models are fitted to the same data set, and the model's rank is evaluated based on which model is getting better goodness of fit statistical measures. Different sets of Rank (1-5) were given to all the models for each of the five G.O.F statistical measures. Then, ranking for derived data sets [$ln(OD/OD_0)$ vs t, $d(OD)/dt$ vs t, $d^2(OD)/dt^2$ vs t, $d(ln(OD/OD_0))/dt$ vs t] from the same data was done in similar fashion. After that, this process was iterated for all available data sets, and then the mean of those ranks of each model was calculated for each of the G.O.F statistical measures and all the different derived sets.

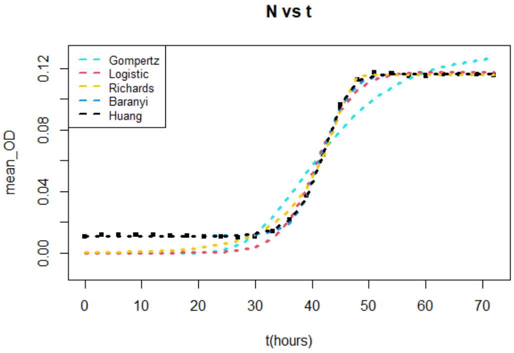
2.6 No. of fits with Best and Worst scores

Similar to Mean ranking calculation, First, all the growth models are fitted to the same set of data and counted which model fits best (smallest AIC, BIC, RSE & largest Adj-R^2, F-stat) and worst (largest AIC, BIC, RSE & smallest Adj-R^2, F-stat) according to each of the five G.O.F statistical measures. This process was also iterated for all available data sets and then calculated total numbers of fits with Best and Worst scores of each model for each of the G.O.F statistical measures and for all the different derived sets.

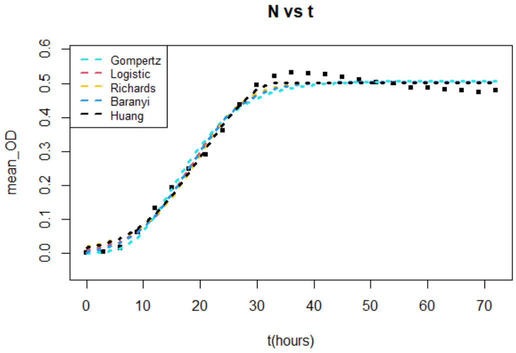
3 Results and Discussion

3.1 N vs t curves

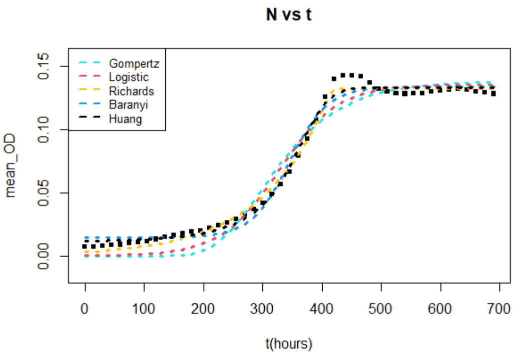
All five models—Gompertz, Logistic, Richards, Barnayi, and Huang—were found to fit the N vs. t scale rather well. However, In comparison to other models, it was seen that the Huang, Richards, and Barnayi models fit the curve slightly better on most of the data-sets Figure 1. The various G-O-F statistical metrics (BIC, AIC, RSE, AdjRsq, F-stats), which account for the number of parameters present in each model, were used to compare the goodness-of-fit Table 1. The lowest mean rank of BIC, AIC, RSE, and AdjRsq corresponded to Huang, followed by Richards, while model Gompertz ranked Highest. With slight exceptions, Richards corresponded to the lowest F-stats Rank followed by logistic, with the same Gompertz ranking Highest. Overall linear average of all the G-O-F statistical measures Rank mean followed the same trend, Huang being lowest and Gompertz Highest Table 1. It was found that for a significant number of curves (57, 62, 81, 81 out of 130), the model Huang had the lowest BIC, AIC, RSE, and highest AdjRsq, respectively. In contrast, model Gompertz had the highest BIC, AIC, RSE, and lowest AdjRsq, F-stats for a significant number of OD vs. t curves (98, 98, 98, 98, 88 out of 130). one exception was in F-stats. It was highest in a significant number of curves (48 of 130) in Richards model Table 1.



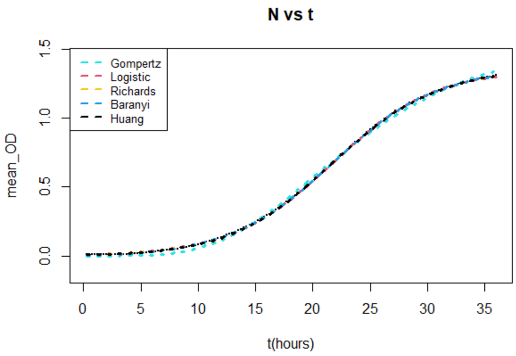
(a) Temp 28C strain App4.8



(b) Temp 35C strain App6.4



(c) V106N



(d) with AB AMS4105 Biol-2

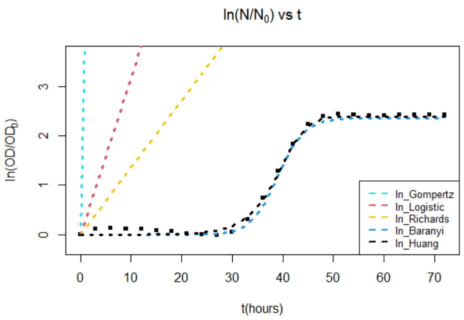
Figure 1: graphs showing all the five model fitted on different OD vs t curves

Table 1: Ranking of models according to mean rank of the different G-O-F statistical measures(for ex: smallest BIC = rank 1, etc.) and the number of curves(total=130) with best/worst fit(best fit means smallest AIC, BIC, RSE but largest AdjRsq and F-stats, and vice-versa for worst fit)on OD vs t scale.

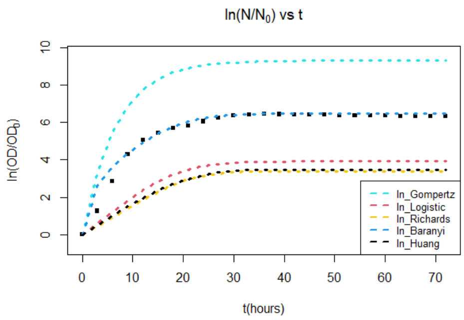
	Gompertz	**Logistic**	**Richards**	**Baranyi**	**Huang**
Mean BIC Rank	4.430769231	3.284615385	2.315384615	2.838461538	2.130769231
No. of Curves with smallest BIC Score	4	18	26	25	57
No. of Curves with largest BIC Score	98	23	1	8	0
Mean AIC Rank	4.469230769	3.376923077	2.330769231	2.838461538	1.984615385
No. of Curves with smallest AIC Score	4	17	23	24	62
No. of Curves with largest AIC Score	98	23	1	8	0
Mean RSE Rank	4.538461538	3.784615385	2.3	2.761538462	1.615384615
No. of Curves with smallest RSE Score	1	2	22	24	81
No. of Curves with largest RSE Score	98	25	1	6	0
Mean AdjRsq Rank	4.546153846	3.8	2.284615385	2.769230769	1.6
No. of Curves with largest AdjRsq Score	1	2	22	24	81
No. of Curves with smallest ADjRSq Score	98	25	1	6	0
Mean F-stats Rank	4.030769231	2.5	2.384615385	3.046153846	3.038461538
No. of Curves with largest F-stats Score	16	48	26	18	22
No. of Curves with smallest F-stats Score	88	6	3	15	18
Avg Mean Rank	4.403076923	3.349230769	2.323076923	2.850769231	2.073846154
Avg No. of Curves with Best Score	5.2	17.4	23.8	23	60.6
Avg No. of Curves with worst Score	96	20.4	1.4	8.6	3.6

This suggests Huang has a substantial edge over other models Barnayi, Logistic, and especially Gompertz in N vs t linear scale. Richard's model seconded Huang and outperformed all other models. This cannot support completely the argument that Mechanistic models (Huang and Barnayi) perform better fit over Empirical models (Gompertz, Logistic, and Richards) cause, in this case, Richards (empirical model) is outperforming Barnayi(Mechanistic model). Exception in F-stats can be explained by the fact that F-stats is calculated to test the hypothesis - that at least one of the coefficients related to predictor variable estimated after fitting is meaningful or non-zero, F-stats gives a value (generally positive), if it is greater than the critical value for that degree of freedom then it has at least one significant variable. So, when all the models fit very well, as in this case, a direct comparison of F-stats of different models does not do justice as it incorporates the degree of freedom also, which will be less in a less complex model like logistic and Richards.

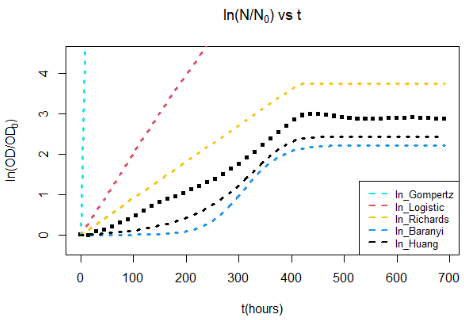
3.2 ln(N/N[0]) vs t curves



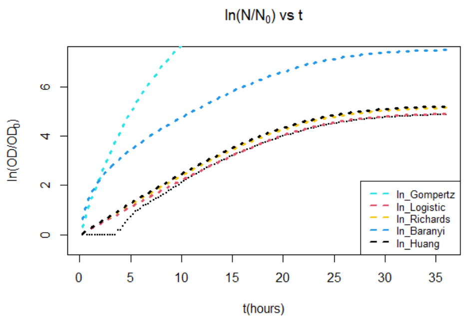
(a) Temp 28C strain App4.8



(b) Temp 35C strain App6.4



(c) V106N



(d) with AB AMS4105 Biol-2

Figure 2: graphs showing all the five model fitted on different ln(OD/OD[0]) vs t curves

When the scale changes from linear to logarithmic (or relative), only models Barnayi and Huang fit the ln(N) vs. t scale somewhat well. However, In comparison to other models, it was seen that the Huang models fit most of the curve better than others Figure 2. Here also, The various G-O-F statistical metrics (BIC, AIC, RSE, AdjRsq, F-stats) were used to compare the goodness-of-fit Table 2. The lowest mean rank of BIC, AIC, RSE, AdjRsq, and F-stats corresponded to Huang, followed by Barnayi, while model Gompertz ranked Highest. Overall linear average of all the G-O-F statistical measures Rank mean followed the same trend, Huang being Lowest and Gompertz Highest Table 2. It was found that for a significant number of curves (48, 48, 48, 47, 48 out of 130), the model Barnayi had the lowest BIC, AIC, RSE, and highest AdjRsq, F-stats, respectively, while model Gompertz had the highest BIC, AIC, RSE and lowest AdjRsq, F-stats

Table 2: Ranking of models according to mean rank of the different G-O-F statistical measures(for ex: smallest BIC = rank 1, etc.) and the number of curves(total=130) with best/worst fit(best fit means smallest AIC, BIC, RSE but largest AdjRsq and F-stats, and vice-versa for worst fit) on ln(OD/OD[0]) vs t scale.

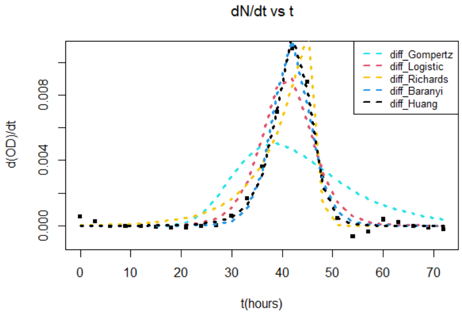
	Gompertz	**Logistic**	**Richards**	**Baranyi**	**Huang**
Mean BIC Rank	4.392307692	3.061538462	2.938461538	2.315384615	2.292307692
No. of Curves with smallest BIC Score	7	23	17	48	35
No. of Curves with largest BIC Score	100	2	11	11	6
Mean AIC Rank	4.407692308	3.076923077	2.938461538	2.307692308	2.269230769
No. of Curves with smallest AIC Score	7	22	17	48	36
No. of Curves with largest AIC Score	100	3	11	10	6
Mean RSE Rank	4.415384615	3.092307692	2.953846154	2.307692308	2.230769231
No. of Curves with smallest RSE Score	7	21	16	48	38
No. of Curves with largest RSE Score	100	3	11	10	6
Mean AdjRsq Rank	4.384615385	3.076923077	2.953846154	2.330769231	2.253846154
No. of Curves with largest AdjRsq Score	5	21	16	47	39
No. of Curves with smallest ADjRSq Score	98	3	11	10	6
Mean F-stats Rank	4.423076923	3.053846154	2.907692308	2.307692308	2.307692308
No. of Curves with largest F-stats Score	6	23	17	48	34
No. of Curves with smallest F-stats Score	105	6	7	8	2
Avg Mean Rank	4.404615385	3.072307692	2.938461538	2.313846154	2.270769231
Avg No. of Curves with Best Score	6.4	22	16.6	47.8	36.4
Avg No. of Curves with worst Score	100.6	3.4	10.2	9.8	5.2

for the relatively high number of ln(OD/OD[0]) vs. t curves (100, 100, 100, 98, 105 out of 130) Table 2.

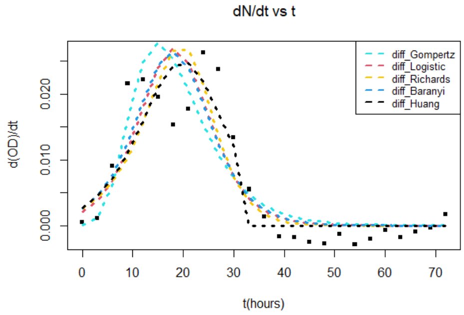
This also suggests similar inference as linear scale paragraph 3.1, i.e, Huang and Barnayi outperform other models like - Richard, Logistic, and especially Gompertz. However, On a Logarithmic scale, the estimation of N_0 (or here $(OD)_0$) plays a critical role for the model to fit accurately, as the population relative to the initial one is being used here for Y-axis. So, this result can help form an inference that Huang and Barnayi model is a better model for estimating the initial population N_0 in most cases. Furthermore, this can also lead to support for the argument- Mechanistic models (Huang and Barnayi) perform better fit over Empirical models (Gompertz, Logistic, and Richards).

3.3 dN/dt vs t curves

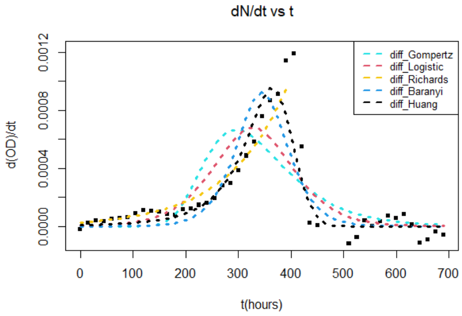
When the scale is changed to derivative (i.e., Absolute growth rate scale), once again, models Barnayi and Huang were found to fit the dN/dt vs. t scale somewhat well in comparison to others. Derivatives here were estimated using cubic smoothing splines techniques to reduce the noise in data. Nevertheless, despite that there is some degree of noise. Still, all the models fit quite well. Figure 3. Here also, the five G-O-F statistical metrics were used to compare the goodness-of-fit Table 3. In this case, the lowest mean rank of BIC, AIC, and F-stats corresponded to Logistic followed by Richards and for RSE and AdjRsq Richards followed by Logistic, while model Gompertz ranked Highest for all the G-O-F metrics except F-stats where Huang was highest. Overall linear average of all the G-O-F statistical measures Rank mean resulted in Logistic being Lowest and Gompertz Highest Table 3. It was found that for a significant number of curves (46, 41, 33, 33, 65 out of 130), the model Logistic had the lowest BIC, AIC, RSE, and highest AdjRsq, F-stats, respectively, while model Gompertz had the highest BIC, AIC, RSE and lowest AdjRsq, F-stats for enough number of d(OD)/dt vs. t curves (73, 74, 78, 78, 62 out of 130) Table 3.



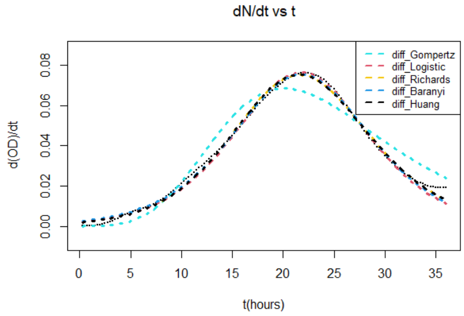
(a) Temp 28C strain App4.8



(b) Temp 35C strain App6.4



(c) V106N



(d) with AB AMS4105 Biol-2

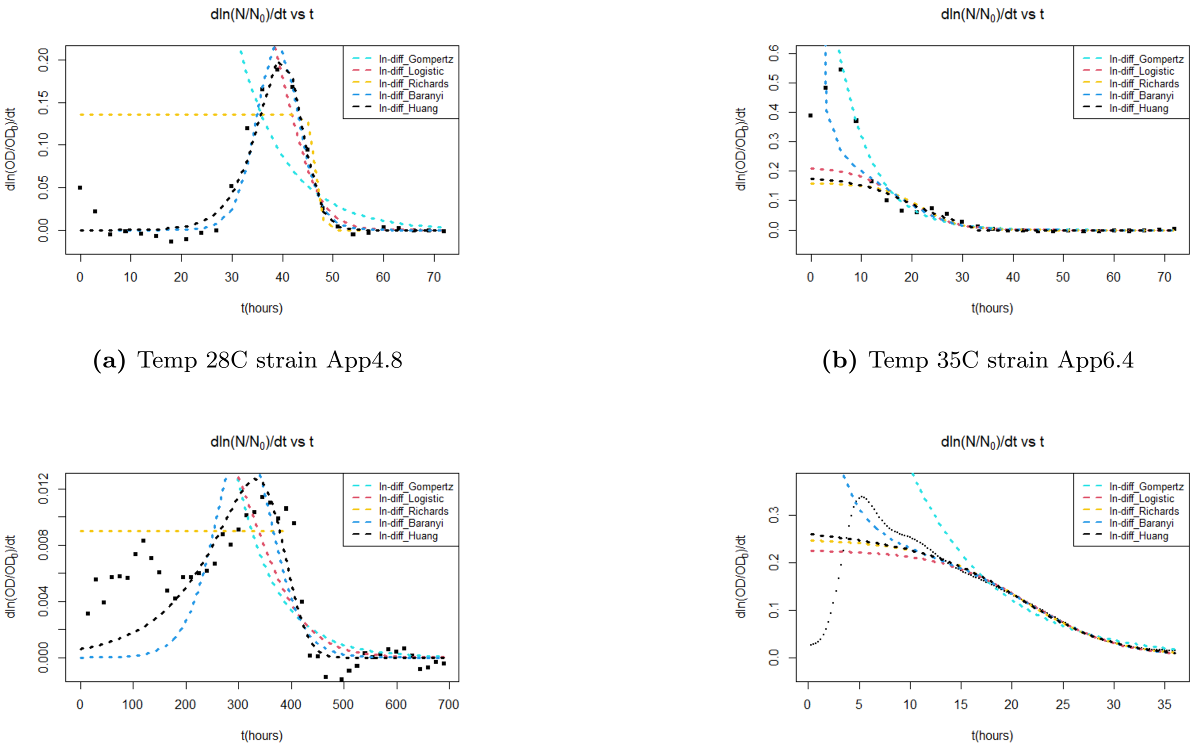
Figure 3: graphs showing all the five model fitted on different d(OD)/dt vs t curves

Table 3: Ranking of models according to mean rank of the different G-O-F statistical measures(for ex: smallest BIC = rank 1, etc.) and the number of curves(total=130) with best/worst fit(best fit means smallest AIC, BIC, RSE but largest AdjRsq and F-stats, and vice-versa for worst fit) on d(OD)/dt vs t scale.

	Gompertz	**Logistic**	**Richards**	**Baranyi**	**Huang**
Mean BIC Rank	3.661538462	2.384615385	2.707692308	2.976923077	3.269230769
No. of Curves with smallest BIC Score	21	46	22	14	27
No. of Curves with largest BIC Score	73	11	8	17	21
Mean AIC Rank	3.661538462	2.523076923	2.646153846	2.946153846	3.223076923
No. of Curves with smallest AIC Score	22	41	24	16	27
No. of Curves with largest AIC Score	74	13	8	16	19
Mean RSE Rank	3.838461538	2.723076923	2.623076923	2.907692308	2.907692308
No. of Curves with smallest RSE Score	16	33	22	21	38
No. of Curves with largest RSE Score	78	14	9	15	14
Mean AdjRsq Rank	3.846153846	2.723076923	2.615384615	2.907692308	2.907692308
No. of Curves with largest AdjRsq Score	15	33	23	21	38
No. of Curves with smallest ADjRSq Score	78	14	9	15	14
Mean F-stats Rank	3.330769231	1.923076923	2.807692308	3.038461538	3.9
No. of Curves with largest F-stats Score	29	65	17	8	11
No. of Curves with smallest F-stats Score	62	6	12	14	36
Avg Mean Rank	3.667692308	2.455384615	2.68	2.955384615	3.241538462
Avg No. of Curves with Best Score	20.6	43.6	21.6	16	28.2
Avg No. of Curves with worst Score	73	11.6	9.2	15.4	20.8

The change in result compared to a linear and logarithmic scale, in addition to that mismatch between the inference drawn from the figure and table, can be explained by the fact that all the five G-O-F account for the number of parameters used in the models and penalize the model for complexity (i.e., greater number of parameters). As it can be seen in Figure 3 all the models are fitting quite well except Gompertz, and the Y-axis value is quite low, so the error will also be of low margin. Therefore Huang, Baranyi, and Richards being more complex got penalized, and Logistic stood out as the best fit model on the derivative scale. This suggests that Empirical models (Logistic and Richards) are effective enough for fitting derivative scale curves.

3.4 d(ln(N/N[0]))/dt vs t curves



(a) Temp 28C strain App4.8

(b) Temp 35C strain App6.4

(c) V106N

(d) with AB AMS4105 Biol-2

Figure 4: graphs showing all the five model fitted on different $d(ln(OD/OD_0))/dt$ vs t curves

This time when the scale is changed to the log of derivative scale (i.e., relative growth rate scale), noise in data is found to be even higher, and despite that, all the models could be seen trying to fit close enough to the curve off-course with some exceptions. Furthermore, among them, models Barnayi and Huang were found to fit somewhat better in comparison to others. Figure 4. Here also, the five G-O-F statistical metrics were used to compare the goodness-of-fit Table 4. This scale also produced similar results like the simple logarithmic result with little swap Barnayi followed Huang resulted in the lowest mean rank of BIC, AIC, RSE, and AdjRsq. In contrast, model Gompertz ranked Highest, But for F-stats, it is the same Huang being lowest, followed by Barnayi. Little swap in lowest and second lowest ranker changed an overall linear average of all the G-O-F statistical measures Rank mean result, i.e, Huang being Lowest and Gompertz Highest Table 4. Here also, it was found that for a significant number of curves (68, 67, , 33, 65 out of 130), the model Logistic had the lowest BIC, AIC, RSE and highest AdjRsq, respectively, whilst model Gompertz had the highest BIC, AIC, RSE, and lowest AdjRsq, F-stats for enough number of curves (95, 95, 95, 95, and 128 out of 130) Table 4.

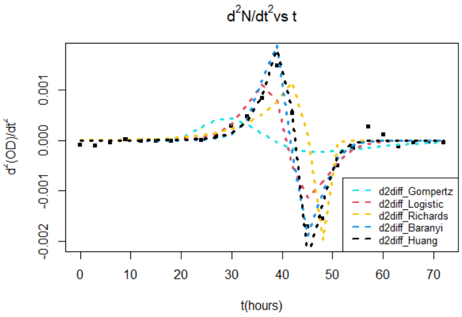
Table 4: Ranking of models according to mean rank of the different G-O-F statistical measures(for ex: smallest BIC = rank 1, etc.) and the number of curves(total=130) with best/worst fit(best fit means smallest AIC, BIC, RSE but largest AdjRsq and F-stats, and vice-versa for worst fit) on ln(d(OD)/dt) vs t scale.

	Gompertz	**Logistic**	**Richards**	**Baranyi**	**Huang**
Mean BIC Rank	4.265625	2.53125	3.1484375	2.21875	2.8359375
No. of Curves with smallest BIC Score	0	44	1	68	15
No. of Curves with largest BIC Score	95	0	10	21	2
Mean AIC Rank	4.28125	2.546875	3.125	2.234375	2.8125
No. of Curves with smallest AIC Score	0	44	1	68	15
No. of Curves with largest AIC Score	95	0	10	22	1
Mean RSE Rank	4.296875	2.59375	3.1328125	2.2421875	2.734375
No. of Curves with smallest RSE Score	0	42	1	67	18
No. of Curves with largest RSE Score	95	0	10	22	1
Mean AdjRsq Rank	4.296875	2.59375	3.1328125	2.2421875	2.734375
No. of Curves with largest AdjRsq Score	0	42	1	67	18
No. of Curves with smallest ADjRSq Score	95	0	10	22	1
Mean F-stats Rank	4.796875	3.796875	2.578125	2.046875	1.78125
No. of Curves with largest F-stats Score	0	0	4	21	103
No. of Curves with smallest F-stats Score	128	0	0	0	0
Avg Mean Rank	4.3875	2.8125	3.0234375	2.196875	2.5796875
Avg No. of Curves with Best Score	0	34.4	1.6	58.2	33.8
Avg No. of Curves with worst Score	101.6	0	8	17.4	1

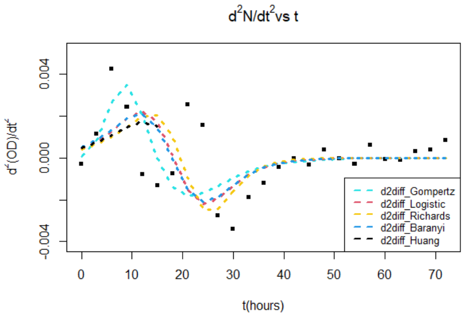
This similar result to linear and especially logarithmic paragraph 3.2 suggests similar inference of Huang and Barnayi outperforming. Also, they are a better model for estimating the initial population N_0 in most cases. Furthermore, Barnayi showing an edge over Huang in this scale can be because of either of two possibilities or both, one similar to derivative scale small error of margin, so less complex model Barnayi (4 parameters than Huang 5 parameters) is getting better G-O-F metrics score or Barnayi performing better on this scale because of it can be better in estimating growth rate (μ) along with initial population (N_0). The first reasoning looks more plausible than the second in this case.

3.5 d2N/dt2 vs t curves

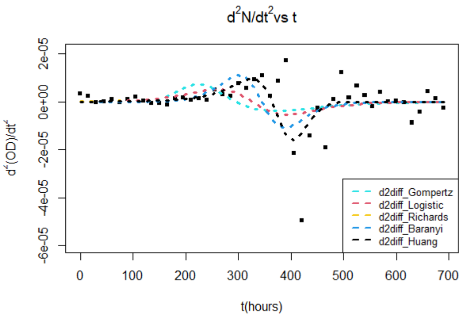
Once again, when the scale is changed to double derivative, noise is relatively high, and despite that, all the models fit well enough to the curve. Among them, model Huang is found to fit somewhat better than others, but the difference is subtle and cannot make an explicit claim just from the graph. Figure 5. one more time, the five various G-O-F statistical metrics were used to compare the goodness-of-fit Table 5. Results from this case are somewhat similar to the derivative one, as the lowest mean rank of BIC, AIC, RSE, AdjRsq, and F-stats corresponded to Logistic, followed by Barnayi, while model Gompertz ranked Highest for RSE, AdjRsq and whereas Huang for BIC, AIC, and F-stats. Overall linear average of all the G-O-F statistical measures Rank mean resulted in Logistic being Lowest and Huang being Highest Table 5. Surprisingly, for a significant number of curves (65, 58, 47 out of 130), the model Logistic was found to have the lowest BIC, AIC, RSE, and (109, 109 out of 130) was found to have the highest AdjRsq, F-stats respectively, whilst model Gompertz had the highest BIC, AIC, RSE for enough number of curves (73, 73, 72 of 130) Table 5 and Logistic had lowest AdjRsq and F-stats for (68, 63 out of 130).



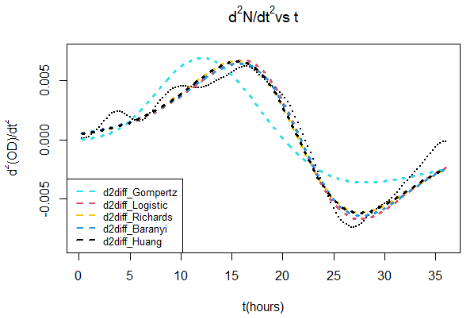
(a) Temp 28C strain App4.8



(b) Temp 35C strain App6.4



(c) V106N



(d) with AB AMS4105 Biol-2

Figure 5: graphs showing all the five model fitted on different $d^2(OD)/dt^2$ vs t curves

Table 5: Ranking of models according to mean rank of the different G-O-F statistical measures(for ex: smallest BIC = rank 1, etc.) and the number of curves(total=130) with best/worst fit(best fit means smallest AIC, BIC, RSE but largest AdjRsq and F-stats, and vice-versa for worst fit) on d2(OD)/dt2 vs t scale.

	Gompertz	**Logistic**	**Richards**	**Baranyi**	**Huang**
Mean BIC Rank	3.230769231	1.953846154	3.146153846	2.946153846	3.723076923
No. of Curves with smallest BIC Score	4	26	62	17	10
No. of Curves with largest BIC Score	73	35	0	12	8
Mean AIC Rank	3.361538462	2.007692308	3.138461538	2.853846154	3.638461538
No. of Curves with smallest AIC Score	4	22	58	18	13
No. of Curves with largest AIC Score	73	36	0	11	8
Mean RSE Rank	3.615384615	2.284615385	3.061538462	2.784615385	3.253846154
No. of Curves with smallest RSE Score	4	20	47	18	19
No. of Curves with largest RSE Score	72	38	1	10	7
Mean AdjRsq Rank	3.630769231	2.307692308	3.046153846	2.784615385	3.230769231
No. of Curves with largest AdjRsq Score	109	7	13	1	0
No. of Curves with smallest ADjRSq Score	1	68	15	21	12
Mean F-stats Rank	3.384615385	1.923076923	3	2.923076923	3.769230769
No. of Curves with largest F-stats Score	109	7	14	0	0
No. of Curves with smallest F-stats Score	0	63	15	16	14
Avg Mean Rank	3.444615385	2.095384615	3.078461538	2.858461538	3.523076923
Avg No. of Curves with Best Score	46	16.4	38.8	10.8	8.4
Avg No. of Curves with worst Score	43.8	48	6.2	14	9.8

There cannot be any strong inference made out of the result obtained in this scale due to a mismatch between the results from different observations and calculations. However, similar to derivative here, also due to low error of margin, Huang, Baranyi, and Richards being more complex, have gotten penalized more. At the same time, Logistic stood out as the best fit in most cases. However, this time Logistic also resulted in max no of the curve with the lowest AdjRsq and F-stats. This can be explained as, if AdjRsq is not in between 0.8 - 1, then only it makes good sense, and comparison has some meaning. But, if the regression model is not linear, the overall mathematical framework for AdjRsq does not function as intended. In non-linear models, underlying assumptions are incorrect. The sum of the explained variance and the error variance IS NOT the total variance, so adjRsq isn't necessarily between 0 and 1. Hence, R-squared doesn't pick the best model every time, specially when noise is high as in this case. AdjRsq will result in some meaningless value despite fitting close enough with a low margin of error. This strengthens the inference that the Empirical model (Logistic) is compelling enough for fitting derivative scale curves.

4 Conclusion

After performing the statistical comparison between sigmoidal kinetic models on several scales, the Baranyi and Huang models performed the best overall. They are better in estimating the initial population N_0 in most cases but with a trade-off complexity. These two models outperformed other with a high margin on *Logarithmic* scale. While still not as good as that obtained with the top two models, the goodness-of-fit obtained with Richards and Logistic can be considered acceptable. Also, they performed very well on *Derivative* scales because the difference in error margin got compensated by less penalty due to less complexity and error difference. Based on many aspects, the Baranyi model demonstrated the best behavior for the growth curves examined. Our results suggest that the Gompertz model, frequently used to describe microbial growth, should be critically reviewed because the Baranyi, Huang, Richards, and Logistic models significantly outperformed the Gompertz in their ability to fit experimental data. Gompertz resulted in worst Goodness-of-fit Scores in most cases on almost all the five scales. Depending on how the model was developed originally, growth models can either be empirical (set out principally to describe) or mechanistic (tries to give a description perhaps with understanding). Typically, a mechanistic model is constructed from a differential equation that connects growth rate (dN/dt) to the organism or population size (N). This mathematical relationship should depict the growth process's mechanism to qualify as a mechanistic model. Our result suggests that most cases of mechanistic models (Huang and Barnayi) perform better than empirical models (Gompertz, Logistic, and Richards). However, empirical models are good enough to describe growth in derivative scales.

Bibliography

Adkar, B. V., Manhart, M., Bhattacharyya, S., Tian, J., Musharbash, M., & Shakhnovich, E. I. (2017). Optimization of lag phase shapes the evolution of a bacterial enzyme. https://www.nature.com/articles/s41559-017-0149

Baranyi, J., Roberts, T., & McClure, P. (1993). A non-autonomous differential equation to modelbacterial growth. *Food Microbiology*, *10*(1), 43–59. https://doi.org/https://doi.org/10.1006/fmic.1993.1005

Baranyi, J., & Roberts, T. A. (1994). A dynamic approach to predicting bacterial growth in food [Special Issue Predictive Modelling]. *International Journal of Food Microbiology*, *23*(3), 277–294. https://doi.org/https://doi.org/10.1016/0168-1605(94)90157-0

Baranyi, J., & Roberts, T. A. (1995). Mathematics of predictive food microbiology. *International Journal of Food Microbiology*, *26*(2), 199–218. https://doi.org/https://doi.org/10.1016/0168-1605(94)00121-L

Buchanan, R., Whiting, R., & Damert, W. (1997). When is simple good enough: A comparison of the gompertz, baranyi, and three-phase linear models for fitting bacterial growth curves. *Food Microbiology*, *14*(4), 313–326. https://doi.org/https://doi.org/10.1006/fmic.1997.0125

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. *Journal of Microbiological Methods*, *43*(3), 183–196. https://doi.org/https://doi.org/10.1016/S0167-7012(00)00219-0

France, J., & Thornley, J. H. M. (1984). *Mathematical models in agriculture : A quantitative approach to problems in agriculture and related sciences / j. france, j.h.m. thornley*. Butterworths London ; Boston.

Gompertz. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. in a letter to francis baily, esq.. f. r. s. amp;c. *Philosophical Transactions of the Royal Society of London*, *115*, 513–583. https://doi.org/10.1098/rstl.1825.0026

Hammer, T. J., Le, E., & Moran, N. A. (2021). Thermal niches of specialized gut symbionts: The case of social bees. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1944), 20201480. https://doi.org/10.1098/rspb.2020.1480

Huang, L. (2011). A new mechanistic growth model for simultaneous determination of lag phase duration and exponential growth rate and a new bělehrádek-type model for evaluating the effect of temperature on growth rate [Predictive Modeling in Foods]. *Food Microbiology*, *28*(4), 770–776. https://doi.org/https://doi.org/10.1016/j.fm.2010.05.019

López, S., Prieto, M., Dijkstra, J., Dhanoa, M. S., & France, J. (2004). Statistical evaluation of mathematical models for microbial growth. *International Journal of Food Microbiology*, *96*, 289–300. https://doi.org/10.1016/j.ijfoodmicro.2004.03.026

López, S., Prieto, M., Dijkstra, J., Dhanoa, M., & France, J. (2004). Statistical evaluation of mathematical models for microbial growth. *International Journal of Food Microbiology*, *96*(3), 289–300. https://doi.org/https://doi.org/10.1016/j.ijfoodmicro.2004.03.026

McMeekin, T., & Ross, T. (2002). Predictive microbiology: Providing a knowledge-based framework for change management [18th International Symposium of the International Committee on Food Microbiology and Hygeine, August 18-23, 2002, Lillehammer Norway. Necessary and Unwanted Bacteria in Food - Microbial Adaption to changing Environments]. *International Journal of Food Microbiology*, *78*(1), 133–153. https://doi.org/https://doi.org/10.1016/S0168-1605(02)00231-3

Model selection and multimodel inference. (2004). Springer New York. https://doi.org/10.1007/b97636

Pearl, R., & Reed, L. J. (1920). On the rate of growth of the population of the united states since 1790 and its mathematical representation. *Proceedings of the National Academy of Sciences*, *6*(6), 275–288. https://doi.org/10.1073/pnas.6.6.275

RICHARDS, F. J. (1959). A Flexible Growth Function for Empirical Use. *Journal of Experimental Botany*, *10*(2), 290–301. https://doi.org/10.1093/jxb/10.2.290

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 and Microbial Technology*, *26*(5), 431–445. https://doi.org/https://doi.org/10.1016/S0141-0229(99)00183-0

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*, 461–464. https://doi.org/10.1214/aos/1176344136

Seber, C. J. W. A. F. (2003). *Nonlinear regression - george*. Wiley. http://as.wiley.com/WileyCDA/WileyTitle/productCd-0471471356.html

Todd, R. T., & Selmecki, A. (2020). Expandable and reversible copy number amplification drives rapid adaptation to antifungal drugs (K. J. Verstrepen & P. J. Wittkopp, Eds.). *eLife*, *9*, e58349. https://doi.org/10.7554/eLife.58349

Vogels, M., Zoeckler, R., Stasiw, D. M., & Cerny, L. C. (1975). P. f. verhulst's "notice sur la loi que la populations suit dans son accroissement" from correspondence mathematique et physique. ghent, vol. x, 1838. *Journal of Biological Physics*, *3*, 183–192.